

# 日本評価研究

## Japanese Journal of Evaluation Studies

Vol. 22, No. 2, September 2022

### 特集：「イノベーション」の形成的評価

特集に寄せて「有効な形成的評価アプローチ法の発展のために」

大島 巖

対人サービスのイノベーションをEBPプログラムに発展させる形成的評価法の開発

－CD-TEP法を用いた実践家・当事者参画型エンパワメント評価の可能性－  
大島 巖 新藤 健太 源 由理子

教育分野における参加型形成的評価の位置付けと意義

米原 あき

社会的インパクト評価の系譜

－マネジメント支援のための評価への進化－

今田 克司

対人サービスにおけるイノベーションをエビデンスに基づく実践プログラムに発展させるための英国の取組み

－児童福祉分野のWhat Works Centreを中心に－

家子 直幸

現場実践を効果モデルに発展させる評価ガイドの開発と実装

－CD-TEP法を活用した事例の考察－

清水 潤子 新藤 健太

実践家参画型エンパワメント評価の基盤を支える「EBP効果モデル」  
技術支援センターの意義と役割

－評価キャパシティ形成に向けた役割を中心に－

新藤 健太 大島 巖

### 研究論文

自治体会計情報のオープンデータ化の促進

－国際基準COFOGによる費目分類に着目して－

生方 裕一 佐藤 亨 川島 宏一

「より良い規制」のための評価システムの条件

村上 裕一

### 研究ノート

業績測定による評価における指標の質改善のための帰納的情報の必要性

池田 葉月

ウランバートル市の大気汚染対策の政策評価

－ロジックモデルの観点から－

オンドラハ バトホヤグ

大学の地域貢献プログラムに質改善を目的とした評価を行うことの意義

岸本 由梨枝

第23回全国大会のご案内

日本評価学会

Japan Evaluation Society

# 日本評価研究

第22巻 第2号 2022年9月

## 目次

### 特集：「イノベーション」の形成的評価

大島 巖

特集に寄せて「有効な形成的評価アプローチ法の発展のために」…………… 1

大島 巖 新藤 健太 源 由理子

対人サービスのイノベーションを EBP プログラムに発展させる形成的評価法の開発

—CD-TEP 法を用いた実践家・当事者参画型エンパワメント評価の可能性—…………… 3

米原 あき

教育分野における参加型形成的評価の位置付けと意義…………… 15

今田 克司

社会的インパクト評価の系譜

—マネジメント支援のための評価への進化—…………… 27

家子 直幸

対人サービスにおけるイノベーションをエビデンスに基づく

実践プログラムに発展させるための英国の取組み

—児童福祉分野の What Works Centre を中心に—…………… 39

清水 潤子 新藤 健太

現場実践を効果モデルに発展させる評価ガイドの開発と実装

—CD-TEP 法を活用した事例の考察—…………… 53

新藤 健太 大島 巖

実践家参画型エンパワメント評価の基盤を支える「EBP 効果モデル」

技術支援センターの意義と役割

—評価キャパシティ形成に向けた役割を中心に—…………… 69

### 研究論文

生方 裕一 佐藤 亨 川島 宏一

自治体会計情報のオープンデータ化の促進

—国際基準 COFOG による費目分類に着目して—…………… 83

村上 裕一

「より良い規制」のための評価システムの条件…………… 97

研究ノート

池田 葉月

業績測定による評価における指標の質改善のための帰納的情報の必要性…………… 111

オンドラハバトホヤグ

ウランバートル市の大気汚染対策の政策評価

—ロジックモデルの観点から—…………… 123

岸本 由梨枝

大学の地域貢献プログラムに質改善を目的とした評価を行うことの意義…………… 137

第 23 回全国大会のご案内…………… 149

【特集に寄せて】

## 有効な形成的評価アプローチ法の発展のために

大島 巖

東北福祉大学

近年世界的にも、既存の制度では対応できない様々な新しい社会課題が拡大し、有効な解決策が模索されている。そのような課題解決に向けて非営利組織など民間セクターの役割への関心も高まっている。これに対して、社会的インパクト評価やエビデンスに基づく政策立案（EBPM）に関心が集められ、課題解決のために新たな発想で生み出された「社会的イノベーション」を、より効果的なプログラムモデルに発展させるために有効な形成的評価の方法論が求められている。

しかしながら現状では、世界的に見ても「効果モデル」を適切に形成・発展させるために有効な形成的評価の方法論は確立していない。そのために、社会課題の解決が十分には進展しない状況がある。

これに対して大島ら（2012、2019）は、実践家・実践現場を基盤にボトムアップ評価（Chen 2015）の方法を用いて、さまざまな利害関係者（stakeholders）が参画し、社会が協働して「EBP効果モデル」を形成・発展させる形成的評価の方法論「CD-TEP法を用いた実践家・当事者参画型エンパワメント評価（PBEE（CD-TEP）」を開発した。また英国では、イノベーションの発展や資金提供のあり方を段階的なものと捉えて、形成的評価の段階や各段階に応じた資金提供のあり方が議論されている。また形成的評価については、教育分野における長い議論の積み重ねが今日でも行われている。

本特集では、社会課題解決のために開発された社会的イノベーションを育て形成・改善させて、より効果的な社会プログラム、さらにはEBP（Evidence-Based Practice）プログラムへと発展させるために求められる、有効な形成的評価のアプローチのあり方を検討する。

その際本特集では、大島ら（2019）が開発し、本学会でも何年かにわたって報告し、会員間の議論を継続して来た「CD-TEP法を用いた実践家・当事者参画型エンパワメント評価（PBEE（CD-TEP）」を主な検討素材に取り上げた。

まず総説の筆頭には、大島らの「対人サービスのイノベーションをEBPプログラムに発展させる形成的評価法の開発－CD-TEP法を用いた実践家・当事者参画型エンパワメント評価の可能性－」を提示した。また、この評価アプローチ法の実践事例としては、清水氏らに「現場実践を効果モデルに発展させる評価ガイドの開発と実装－CD-TEP法を活用した事例の考察－」をご執筆頂いた。さらに、この形成的評価を進める上で不可欠な「EBP効果モデル」技術支援センターの意義と役割を、新藤氏らには「実践家参画型エンパワメント評価の基盤を支える「EBP効果モデル」技術支援センターの意義と役割－評価キャパシティ形成に向けた役割を中心に－」をテーマにまとめて頂いた。

一方、上述したとおり、世界における形成的評価の発展において、教育分野での取り組みは重要である。本特集の総説の一つとして、米原氏には「教育分野における参加型形成的評価の位置付けと意義」をまとめて頂いた。

さらに、「社会的イノベーション」の評価領域では、近年社会的インパクト評価に関心が集められている。形成的評価、総括的評価も共に求められるこの領域における評価をどのように位置づければ良いのか、現在発展途上にある社会的インパクト評価の総説を、今田氏には「社会的インパクト評価の系譜－マネジメント支援のための評価への進化－」としてご執筆頂いた。

また世界的な動向として、英国の国家的なEBPMに関わる形成的評価の取組みが注目されている。家子氏には、英国の実践報告として「対人サービスにおけるイノベーションをエビデンスに基づく実践プログラムに発展させるための英国の取組み－児童福祉分野のWhat Works Centreを中心に－」をまとめて頂いた。

なお本特集は、2019年12月に高知大学で行われた日本評価学会第20回全国大会における共通論題1「イノベーションの形成的評価」の成果と議論を、さらに深化・発展させたものである。本特集執筆者も、その共通論題1の報告者、討論者、そして当日会場での発言者から構成されている。その後も、このメンバーでは複数回、特集テーマに関する意見交換の場を持ち、このテーマに関する共通認識を深めて来た。

「形成的評価」の概念・用語は、評価学関係者の間では良く知られたものであるが、前述したとおり、その有効な評価方法論に関する議論は必ずしも十分に深められていない。

この特集号が、いま社会的に必要とされる「社会的イノベーション」の形成・発展を促す有効な形成的評価の方法論として社会的な認知が広がり、「形成的評価」の有効な方法論に対する議論が活性化することを、心から願っている。

## 参考文献

Chen, HT (2015) .*Practical program evaluation*. SAGE.

大島巖、ほか (2012) 「CD-TEP | 円環的対話型評価アプローチ法実施ガイド」、平成22年度文部科学省・科学研究費補助金基盤研究 (A) 『プログラム評価理論・方法論を用いた効果的な福祉実践モデル構築へのアプローチ法開発報告書』 (主任研究者：大島巖)、<http://cd-tep.com/> (2019.10.31取得)

大島巖・源由理子・山野則子、ほか (2019) 『実践家参画型エンパワメント評価の理論と方法—CD-TEP法：協働によるEBP効果モデルの構築』、日本評論社

## 【総説】

# 対人サービスのイノベーションを EBP プログラムに 発展させる形成的評価法の開発 － CD-TEP 法を用いた実践家・当事者参画型エンパワメント評価の可能性－

大島 巖  
東北福祉大学  
i-oshima@nifty.com

新藤 健太  
日本社会事業大学  
shindo@jcsu.ac.jp

源 由理子  
明治大学  
yminamoto@mvd.biglobe.ne.jp

## 要 約

こんにち対人サービスの領域では、既存サービスでは対応できない多様な社会課題が拡大し解決方策が模索されている。これに対して課題解決のために新たな発想で生み出された社会的イノベーションの「効果モデル」を適切に形成・発展させて、エビデンスに基づく実践プログラム（EBP）へと成長させる形成的評価法が求められている。

これに対して著者らは、実践現場を基盤に協働型評価、ボトムアップ評価を用いて、実践家と当事者が評価活動に積極的・主体的に参画し、協働して「EBP効果モデル」を形成・発展させる形成的評価法「CD-TEP法を用いた実践家・当事者参画型エンパワメント評価」を開発した。

本稿では、この評価法の開発理念と開発プロセスを示し、その結果開発された評価法の概要と基本枠組みを提示した。その上で他の評価法との比較からその意義と特色、有用性、適用範囲と課題を明らかにし、今後の発展可能性を考察した。

## キーワード

形成的評価、協働型評価、実践家・当事者参画型エンパワメント評価、ボトムアップ評価、エビデンスに基づく実践プログラム（EBP）

## 1. はじめに

こんにち世界的なレベルで、社会政策のさまざまな領域でエビデンスに基づく政策立案（EBPM：Evidence-Based Policy Making）を重視する動きが急速に進展している（古矢 2017）。同時に、社会課題の解決のために、社会的インパクト（社会的な成果）に注目して、課題解決に有効な社会プログラム構築を科学的なエビデンスに基づいて進める評価が公共・民間の両領域で

推進されている（Boazら 2019; 大島 2019）。

このような中、日本の対人サービスの領域では、既存サービスでは対応できない社会課題が拡大している。これに対して、新たな発想のもと革新性の高い課題解決方法として生み出された社会的イノベーションの「効果モデル」（効果をもたらす社会プログラムのモデル）を適切に育て上げ、より効果的なエビデンスに基づく実践（EBP：Evidence-Based Practice）プログラムへと形成・発展させる有効な形成的評価法が

社会的に求められている (Boazら2019; 大島ら2019)。

しかしながら現在、世界的に見ても適切で有効な形成的評価の方法論は確立していない。そのために、社会課題の解決が十分に進展しない状況があると考えられる。これに対して大島ら (2012; 2019) は、参加型・協働型評価 (Rossiら 2004) や、実践現場を基盤により効果的なプログラムモデルを求めるボトムアップ・アプローチによる評価 (ボトムアップ評価) (Chen 2015: 400-403) の方法を用いて、さまざまな利害関係者 (stakeholders)、中でも日常的にプログラムに接する実践家とサービス受益者 (以下、当事者) が積極的・主体的に評価活動に参画・協働して、EBPなどの「効果モデル」(以下「EBP効果モデル」、単に「効果モデル」とも呼ぶ) を形成・発展させる形成的評価の方法論「CD-TEP法を用いた実践家・当事者参画型エンパワメント評価」を開発した (大島ら 2019)。

ここで私たちが提案するCD-TEP法とは、「プログラム理論 (Theory)」と「エビデンス (Evidence)」と「実践 (Practice)」との「円環的対話 (CD: Circular Dialogue)」を通じて、社会プログラムに関わる関係者が協働で、EBPなどの「効果モデル」の構築を目指す評価手法である (大島ら 2012; 2019)。関係者間の円環的な対話 (CD) を可能にするために、「効果モデル」をプログラム理論や「効果的援助要素」など5アイテム (後述) で可視化するとともに、「効果モデル」の形成・発展のための評価方法・評価計画も「CD-TEP改善ステップ」(後述) で同様に可視化する。それによって、実践家・当事者およびその他関係者たち (実践家等と呼ぶ) が協働して評価活動に参画することを可能にする体系的な評価手法である。

本論文では、この新しい形成的評価法の開発理念と開発プロセスを示し、その結果開発された評価法の概要、基本的な枠組みを整理して提示する。その上で、他の評価アプローチ法との比較からみた意義と特色、有用性、適用範囲と課題について検討し、今後の発展可能性を考察する。

## 2. 新しい形成的評価法の理念・特色、開発方法

新しい形成的評価法の開発に当り私たちが目指すもの、理念・特色をまず明らかにする。その上でその評価法の開発方法を提示する。

### (1) 理念・特色、目指すもの

私たちが目指すもの、理念・特色は、以下の6点に整理できる。

#### ①EBPプログラム志向: 「効果モデル」改善の方向性、プロセスの明示

まず新しい形成的評価法は、「EBP効果モデル」の形成・発展を目指している。そのために「成果への影響」を向上させること、すなわちエビデンスレベル (津谷 1999) や効果量の向上 (内的妥当性の向上 (大島ら 2020: 64-65))、および「効果モデル」の社会実装の推進 (外的妥当性向上を含む (大島ら 2020: 64-65)) を目指すことを明確に位置づけた。そのために「効果モデル」が形成・発展するプロセスを「評価ステージ」として具体的に示した。

#### ②評価5階層の体系的活用

形成的評価では、5つの評価階層 (Rossiら 2004: 80) を体系的かつ戦略的に活用する。特にプログラム理論評価とプロセス評価を体系的に活用し、同時にアウトカム評価と関連付けて、「成果への影響」を最大化する有効な評価方法を提示する。

#### ③ボトムアップ評価: 実践現場から「効果モデル」の形成・発展を目指す

実践現場のニーズに根ざして、実施可能な評価を積み重ねて「効果モデル」の形成・発展を目指す。そのために日常的にプログラムに接し、それを熟知する実践家等が参画し、「効果モデル」に発展させるボトムアップ評価の方法を取り入れる (Chen 2015: 400-403)。

④可視化・共有化：協働型評価に必要な可視化・共有化を行う

協働型評価を適切に実施するために、評価活動に参画する実践家等に対して、評価に必要な情報を可能なかぎり可視化して提供し、共有する必要がある。

⑤エンパワメント評価：評価活動に参画してエンパワーされること

日常的にプログラムに接する実践家等が評価活動に参画し、より優れた成果をもたらす「効果モデル」を形成し、課題解決を前進させることにより、関与した実践家等はそれぞれの自己効力感を向上させ、エンパワーされる。同時にその活動によって実践家等の評価キャパシティ(Labinら 2012)も向上する。

以上は「エンパワメント評価」の特徴であり、定義の一部でもある(Wandersman 2005: 27-28)。このことを踏まえて、私たちはこの評価方法を「実践家・当事者参画型エンパワメント評価(Practitioner & peer Based Empowerment Evaluation: PBEE)と命名した。またPBEEを具体的に進める「評価手法」として、私たちが開発したCD-TEP法を活用する(以下PBEE(CD-TEP)評価法と呼ぶ)。

⑥技術支援センター：実践現場をサポートするシステムと一体的に体系化

実践現場に基盤を置く評価活動を、安定的にかつ効果的に行うためには、実践現場をサポートする中間支援組織である「技術支援センター(Technical Assistance Center)」(Blasé 2009; Goughら 2018)、技術移転支援専門組織あるいは人材(purveyors)の存在は不可欠である(Fixsen 2005)。PBEE(CD-TEP)評価法は、実践現場をサポートする技術支援システムと一体的に体系化する必要がある。

(2) 新しい形成的評価法の開発方法とその発展プロセス

①研究全体の発展プロセスの概要

著者らの研究チームは、2007年度以来文部科学

省科学研究費補助金(基盤研究(A)を継続3期受けて、新しい対人サービスの実践プログラム(社会イノベーション)を、エビデンスに基づく効果モデルに形成・発展させる評価アプローチ法を発展させて来た(大島ら 2011; 2015; 2019; 2022)。

検討した社会プログラムは、主に福祉実践プログラムであり、各時期とも10前後のプログラムを取り上げた。各プログラム班とも共通フォーマットに基づき各々の形成・発展ステージから次のステージへ形成・発展する取組みを進めた。評価アプローチ法の発展については、各分野(高齢、子ども・家庭、障害、精神保健福祉、地域福祉等)の関係者が、年に数回、全体会を持ち集中的な議論を重ねながら合意を得た。

②第1期研究(2007-10年度): CD-TEP評価法の生成

「効果的な福祉実践モデル構築への評価アプローチ法」を開発することを目標とした取組みである(大島ら 2011)。その成果はCD-TEP評価アプローチ法としてまとめられた。実践プログラムに関する「プログラム理論(T)」と「エビデンス(E)」「実践(P)」を相互に円環的に関連させ、課題プロセスを前進させるために必要な共通基盤6項目を抽出した[(1)測定可能なプログラムゴールの設定と共有、(2)合意できるプログラム理論の形成、(3)効果的援助要素の作成と共有、(4)チェックボックス方式による効果的援助要素の記述と測定、(5)効果的援助要素チェックボックスに基づく実施マニュアルの構築、(6)プログラムゴールとなるアウトカム指標と効果的援助要素の関連性]。この6項目を整理し、「効果モデル」構成要素として抽出したのが、後述する「効果モデル5アイテム」である。

③第2期研究(2011-14年度): CD-TEP法改善ステップの生成

「実践家参画型福祉プログラム評価の方法論および評価教育法」の開発を旨とした取組みであり、CD-TEP改善ステップがその成果物としてまとめられた(大島ら 2015)(後述)。

④第3期研究 (2015-21年度) : 有効なEBP技術支援センターガイドライン構築

「実践家参画型エンパワメント評価の実施基盤を支える有効なEBP技術支援センターガイドライン構築」を目的とした取組みである (大島ら 2022)。各プログラム班に共通するEBP技術支援センター (EBP-TAC) の機能と役割に関する指針を9領域にまとめた (大島ら 2022 : 9-15)。

3. 開発された「CD-TEP法を用いた実践家・当事者参画型エンパワメント評価」の実施方法

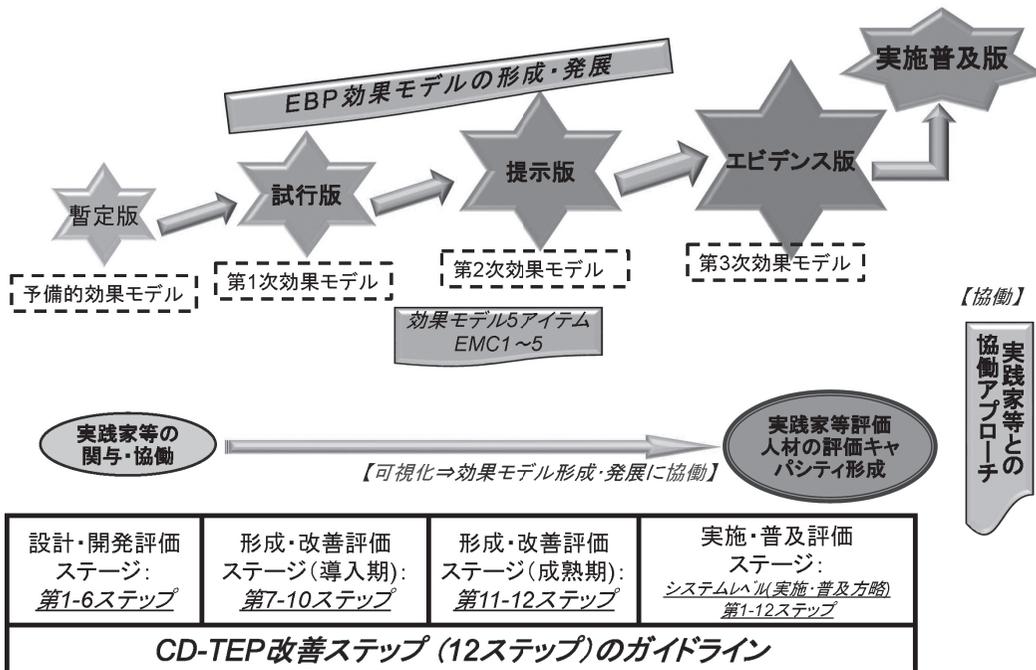
以下では、開発された形成的評価法の概要と基本的な枠組みを提示する。加えて前節1項で示した「理念・特色、目ざすもの」(以下「理念・特色」) との関係にも言及する。

(1) CD-TEP法の概要

PBEE (CD-TEP) 評価法では、形成的評価の実施によって「効果モデル」の「版」が更新されて、効果性と普遍性がより優れた「効果モデル」に形成・発展することを想定する (図1参照 ; 暫定版→試行版→提示版→エビデンス版→実施普及版)。CD-TEP改善ステップのガイドラインに基づいてプロセスを前進させる。

形成的評価の実施は、実践家や当事者・家族、市民等 (実践家等) の参画を得て、ボトムアップ評価 (Chen 2015 : 400-403) の方法を用いて進める。実践家等が「効果モデル」とその改善方法を共有することにより、実践現場の創意工夫や実践上の知見・アイデア、実践知が蓄積され、より良い「効果モデル」が形成・発展することを想定する。そのため、《1》「効果モデル」の可視化については、効果モデル5アイテム (後述) を設定する。また《2》形成的評価の方法については、CD-TEP改善ステップのガイドラインで可

図1 CD-TEP改善ステップの模式図



(出所) 大島ら (2019)

視化・共有化する（後述、図1参照）（大島ら2019）。

## (2) 「効果モデル」形成・発展ステージ

「効果モデル」の形成・発展の目的に対して、PBEE（CD-TEP）評価法は、その評価プロセス全体を通じて、「効果モデル」の設計と開発、形成と改善、および社会の中での実施・普及に必要な情報を提供する役割を担う。「効果モデル」の形成・発展プロセスには、3つの評価課題と、各課題に対応した4つの評価ステージが含まれる。

3つの評価課題とは、第1に「効果モデル」を設計・開発する課題、第2にその継続的改善と形成の課題、第3に実施・普及の課題である。

「CD-TEP改善ステップ」（図1）では、Chen（2015：36-41）を参考にして「形成・改善評価ステージ」を「導入期」と「成熟期」に分け、以下の4段階ステージ（効果モデル形成・発展ステージ）を設定した。

- ・第Iステージ：設計・開発評価ステージ
- ・第II-1ステージ：形成・改善評価ステージ（導入期）

・第II-2ステージ：形成・改善評価ステージ（成熟期）

・第IIIステージ：実施・普及評価ステージ

この「CD-TEP改善ステップ」および「効果モデル形成・発展ステージ」は、前節「理念・特色」の①EBPプログラム志向に対応する。同時に、前節「理念・特色」④可視化・共有化のうち、評価方法・評価計画の可視化・共有化にも対応する。

## (3) CD-TEP法による可視化・共有化

CD-TEP法では「効果モデル」の可視化・共有化と、「効果モデル」形成・発展のための評価方法の可視化・共有化を重視する。実践家・当事者・家族・市民など関係者が参画し、協働・共創により社会課題の解決を旨とする上では不可欠からである（前節「理念・特色」の④）。

まず、「効果モデル」の可視化・共有化では、

【1】「効果モデル」が達成を目指している目標やゴール、と【2】ゴール・目標を達成するため行われるプログラムの実施方法（提供組織のあり方、サービス提供方法）の可視化・共有化を重

表1 効果モデル5アイテム

### 「効果モデル」を構成する要素 (Effective Model Component 1-5) を、次の5点に整理

- EMC1) 支援ゴールとインパクト理論: 効果モデルのゴールと、その達成過程を示すプログラム理論のインパクト理論
- EMC2) プロセス理論: プログラムゴールを実現するために有効なプログラム設計図に当たるプロセス理論(サービス利用計画, 組織計画)
- EMC3) 効果的援助要素リスト: チェックボックス形式で記述した効果的支援のエッセンスのリスト。EMC2)プロセス理論の構成要素
- EMC4) 評価ツール: EMC3)効果的援助要素に基づくモデル適合度尺度(フィデリティ評価尺度)、およびEMC1)に基づいて、プログラムアウトカムを随時測定・モニターするための評価のツール
- EMC5) 実施マニュアル: 以上EMC1-4)の内容を具体的に記載した効果モデル実施マニュアルと評価マニュアルから構成される

(出所) 大島ら (2019)

視する。

これに対して、「効果モデル」の構成要素を抽出したのが、「効果モデル5アイテム」である（表1）。【1】の「効果モデル」の目標やゴールについては、評価階層ではプログラム理論評価レベルの「EMC1）支援ゴールとインパクト理論」が対応する。同時にアウトカム評価レベルの「EMC4）評価ツール」のアウトカム指標・尺度が活用される。【2】のプログラム実施方法については、評価階層ではプログラム理論評価レベルの「EMC2）プロセス理論」と、プロセス評価レベルの「EMC3）効果的援助要素リスト」、「EMC4）評価ツール」のフィデリティ尺度、「EMC5）実施マニュアル」が対応する。

「効果モデル5アイテム」に関しては、前節「理念・特色」との関係では、「②評価5階層の体系的活用」が行われる。評価5階層のうち、特にプログラム理論評価レベル、プロセス評価レベルの位置づけが、5アイテムにおいて重視されている。

また「EMC3）効果的援助要素リスト」は、前節「理念・特色」③ボトムアップ評価の側面から特に重要である。「EMC3）効果的援助要素リスト」で用いられるチェックボックス項目の多くは、実践現場の創意工夫や優れた実践知などが反映されている。

なお前述のとおり【3】「EBP効果モデル」の形成・発展のための評価方法・評価計画法も「CD-TEP改善ステップ」で可視化した（図1）。

#### (4) 「評価計画の設計」に寄与する「CD-TEP改善ステップ」の提示

CD-TEP改善ステップは、解決が求められる社会課題に対して、より優れた「効果モデル」を形成・発展させるために、PBEE（CD-TEP）評価法のモデル的評価実施プロセスを提示した指針である。各形成・発展ステージにおいて、どのような評価活動を行い、次のステージ（版）の「効果モデル」に形成・改善するのか、その評価活動の枠組みとガイドラインを4つの形成・発展ステージごとに提示した。

CD-TEP改善ステップに基づく評価方法・評価計画は、対象となる「効果モデル」の類型（試

行版か、エビデンス版かなど）が明らかになれば、次の評価課題が提示できる特徴をもつ。このため必要な評価計画の設計が、各ステージのガイドラインに依拠して容易に行うことが可能になる。

これは前節「理念・特色」④評価に必要な可視化・共有化が対応する。

#### (5) 「評価の場」と効果モデル改訂の方法

「CD-TEP改善ステップ」では、実践現場と評価活動の相互交流の場（「評価の場」）として、実践家等参画型ワークショップ、GP事例調査、試行評価プロジェクトの評価訪問など7場面の「評価の場」を設定した。

「評価の場」では、実践家等から多くの「実践知（P）」がインプットされる。その結果を踏まえて、質的・量的評価データの収集と分析という科学的な評価活動の知見・エビデンス（E）が生成され、合わせて実践家等参画型ワークショップなどで効果モデル改訂などの議論が行われる。

#### (6) 実践家等の評価キャパシティ形成方法

エンパワメント評価では、評価目的の達成は実践家等関係者の評価キャパシティの向上によりもたらされる（Wandersman 2005: 27-28）。前項の「評価の場」は、その場への関与を進めることで「評価キャパシティ形成の場」になることを意図した戦略的なアプローチが必要になる。

実践家等の評価キャパシティ形成を含めて、PBEE（CD-TEP）評価法を有効に進めるためには中間支援組織としての技術支援センターの役割が重視される（前節「理念・特色」⑥）。

## 4. 他の評価法との比較からみた意義と特色、有用性

以下では、開発されたこの新しい形成的評価法が、社会課題解決のため導入された社会的イノベーションの形成・発展に有効な方法であるか、他の評価アプローチ法との比較から検討したい。

その際、まずは第2節1項の「理念・特色」で示した6項目の中から、形成的評価に期待される役割として重要な次の5機能を整理し検討する(第2節1項「理念・特色」の③④は、③に統合)。

すなわち、①プログラム改善の方向性：EBPプログラム志向、②評価5階層の体系的活用、③ボトムアップ評価、④エンパワメント評価、⑤実践現場をサポートする技術支援システム構築、である。

### (1) プログラム改善の方向性：EBPプログラム志向

まずPBEE (CD-TEP) 評価法は、CD-TEP改善ステップ(図1)に見るように、「EBP効果モデル」の形成・発展を目ざしている。「EBP効果モデル」の効果に対する内的妥当性を高めるとともに、社会実装を進めて、さまざまな地域・社会・文化においても普遍的に有効性をもつ外的妥当性の高い、対人サービス領域の効果的モデル構築を可能とする枠組みを用意する。

これに対して、形成的評価の一般的な定義は、「プログラムの改良を導くための情報提供を意図した評価(活動)」(Rossiら 2004:34)である。「プログラムの改良」には、もちろん内的妥当性、外的妥当性が担保された普遍性の高い「EBP効果モデル」への形成・発展は含まれるが、それがすべてではない。

特定の地域や組織での適用と有効性の向上、地域・組織でのプログラムマネジメントの向上も含まれるであろうし、さらには個別のプログラム利用者ごとの適用と有効性の向上も含まれる場合がある。たとえば、本誌特集で米原(2022)が指摘するように、教育分野の参加型形成的評価では、特定の地域や学校における教育活動の改善を目ざす評価が重視されるだろう。

他方、「EBP効果モデル」の形成・発展を目ざす形成的評価は、近年その有効な評価法が注目され、より良い方法が模索されている。本誌特集で家子(2019; 2022)が取上げた英国政府主導で設置された知識仲介機関What Works Centre (WWC)の1つであるEIF (Early Intervention Foundation)の「評価成功の10ステップ」(Asmussenら 2019)や、政府基金による外郭団

体で設立されたイノベーション推進組織Nestaのエビデンス尺度(Standards of Evidence)の5レベルには形成的評価を活用してエビデンスレベルを向上させ社会実装する仕組みが盛り込まれる(Puttickら 2013)。

また医学領域のコクラン共同計画と並ぶ、社会サービス領域の国際的データベース、キャンベル共同計画(Campbell Collaboration: C2)では、C2のCEOであるWhite(2019)が、エビデンスに基づくプロジェクトサイクルをまとめて提言を行い、サイクルの最初に位置する「形成的評価」の適切な活用を促す。しかしながら、その具体的な方法は明確に示されていない。さらにC2のシステムティックレビュー(SR)には、形成的評価の知見は反映されていない(Littellら 2017)。

ところで、「EBP効果モデル」の形成・発展という方向性をもつ形成的評価では、「評価ステージ」のプロセスの明確化が不可欠である。PBEE (CD-TEP) 評価法で使用する「効果モデル形成・発展ステージ」は、「CD-TEP改善ステップ」の各ステージに記載される評価ガイドラインとともに、形成的評価のプロセスを明示する。

「CD-TEP改善ステップ」は、Chen(2015)のプログラムステージとCORE(2009)プログラムライフサイクルを参考に作成した。これらは、それぞれに有用な「評価ステージ」であるが、Chen(2015)のプログラムステージには、実施・普及段階の評価が位置付いていないこと、CORE(2009)のプログラムステージにはライフサイクルに対応する形成的評価の方法が明確に記載されていない点でPBEE (CD-TEP) 評価法の「評価ステージ」とは位置づけが異なる。

### (2) 評価5階層の体系的活用

前述のとおり、形成的評価の主要な特徴は、5つの評価階層(Rossiら 2004:80)を体系的に活用することにある。中でもプログラム理論評価とプロセス評価を体系的に活用して、さらにはアウトカム・インパクト評価と関連付けて、質的・量的に実証的に分析することが重要である。

PBEE (CD-TEP) 評価法は、プログラム理論評価やアウトカムに結びつくプロセス要素の分析に基づく「効果モデル5アイテム」の活用を中核

にして、評価5階層の体系的活用を具体的な指針とする特長をもつ。

一方、EBPの国際的データベースを提供するキャンベル共同計画（C2）には形成的評価の視点が乏しく、形成的評価の知見は反映されていないことは先ほど触れた通りである。他方、EIFの「評価成功の10ステップ」（Asmussenら 2019）や、Nesta（2018）のエビデンス尺度5レベルについても、評価5階層の体系的・戦略的活用は明確に記載されていない。

### (3) ボトムアップ評価：協働型評価に必要な可視化・共有化

PBEE（CD-TEP）評価法は、実践家等参画型で進めるボトムアップ評価（Chen 2015：400-403）の特徴がある。日常的にプログラムに接し、それを熟知する実践家等が積極的に評価活動に参画し、実践現場の創意工夫や優れた実践知などを十分に反映して、より良い課題解決を可能とする優れた「効果モデル」を生み出すことを期待する。さらには「効果モデル」に関わる関係者相互の間で、また実践家等と評価研究者の間で、日常的に評価情報を共有し、その内容を相互に検証して日々改善に役立てることが、実践現場において評価文化を形成して、職場を「学習する組織」（Senge 2006）に変革する基盤になる。

ボトムアップ評価は、実践現場の実践家等が評価活動に主体的に参画することから、自ずから協働型評価になる。その取組みを進めるに当たって、「効果モデル」および「評価方法・評価計画」の可視化・共有化は不可欠である。しかし「ボトムアップアプローチ」の概念を生み出したChen（2015）の著書には、この可視化・共有化の具体的な方法が十分には記載されていない。

### (4) エンパワメント評価の導入

「EBP効果モデル」の形成・発展という方向性をもつ形成的評価において、日常的にプログラムに接する実践家等参画型のエンパワメント評価を活用した点がPBEE（CD-TEP）評価法のユニークな点であり、この評価法の特長と言える。

### (5) 実践現場をサポートする技術支援システム構築

以上のとおりPBEE（CD-TEP）評価法は、実践に根ざした有効な形成的評価の方法論となる可能性がある。

しかし一方で、この評価法は課題解決に意欲がある資質の高い実践家等の存在がある程度前提になる。また実践家等に必要な評価キャパシティが備わって適切な形成的評価の実施が可能になる側面がある。

このような状況を作り出すためにも、実践現場を日常的にサポートする中間支援組織、技術支援センターの存在は不可欠である（Blase 2009; Goughら 2018; Fixsen 2005）。前述した英国の知識仲介機関What Works Centre（WWC）はそのような役割が期待される組織の1つである。

近年日本でも、実践現場と研究・専門機関や行政等の中間に立って、両者をつなぎ、コミュニケーションを取りながら、実践現場に必要な技術支援等を提供する拠点組織（中間支援組織）が注目されている（木村 2016）。有効なPBEE（CD-TEP）評価法実施のためにもその体制整備が必要になる。

## 5. PBEE（CD-TEP）評価法の適用範囲について

以上検討してきたとおり、PBEE（CD-TEP）評価法は、対人サービスの新しい実践プログラム（社会イノベーション）を、より効果的なエビデンスに基づく実践プログラムへと形成・発展させる体系的な形成的評価法として一定の有用性があると考えられる。

一方で、この評価法が適用可能となる範囲や課題も明らかになったと考える。

まず第1に、形成的評価が目ざす「改善の方向性」についてである。PBEE（CD-TEP）評価法は、「EBP効果モデル」の形成・発展という明確な方向性をもつ。これに対して、社会課題が地域ごとに多様で、その解決の主体や使う社会資源も異なる社会介入では、実践家参画型協働評価の有効性は認められるも（源 2014）、PBEE

(CD-TEP) 評価法の適用は必ずしもなじまない場合がある。

また第2に、評価対象となる社会プログラムが扱う社会課題の特性についてである。PBEE (CD-TEP) 評価法は、保健・医療・福祉・労働などの対人サービスに関わる社会課題であり、しかもそれが比較的深刻で、地域的にも幅広く広がっている場合に有用と考える。たとえば、子どもの貧困、遷延化するひきこもり・8050問題、ヤングケアラー、虐待、脱施設化、違法薬物の再犯防止などの社会課題である。

他方で、対人サービス以外の領域（道路・河川などのインフラ整備など）や、文化・芸術、社会教育などの領域では他の評価方法がふさわしい場合があるだろう。

第3に社会課題の緊急的対応の必要性である。PBEE (CD-TEP) 評価法は、比較的十分に時間をかけながら、EBPなど「効果モデル」を構築することを意図した評価法である。しかし災害やパンデミックなど即応性が問われる社会課題に対する緊急対応の場合は対応が難しい。発展的評価 (Patton 2021) などの他の評価方法がふさわしい場合があると考えられる。

## 6. 今後の課題と展望

PBEE (CD-TEP) 評価法で用いるボトムアップ評価、実践家等参画型協働評価は、いずれも専門の評価者や評価研究者ではない人材が、主体的に形成的評価に関わり、評価専門家とも協働して評価を実施するアプローチである。PBEE (CD-TEP) 評価法では、専門の評価者や評価研究者も実践家等と一緒に評価研究チームに加わることを想定する。しかしながら実践家等が、評価研究チームに関わり、評価に関わる一定の役割を担うのは、必ずしも容易ではない。著者らの15年近い共同研究の実績に基づく指針はあるものの (大島ら 2019)、今後も引き続き、実践家等および評価者・評価研究者が評価チームにおいて、それぞれの役割を分担するための経験知を蓄積することが必要になるだろう。

実践現場に基盤を置く評価活動を、安定的か

つ効果的に行うために、実践現場をサポートする中間支援組織や技術支援センターなど、技術支援体制の整備は不可欠である。

これに対して、アメリカの教育技術支援センター (IES 2011) や、イギリスのWWC (Goughら 2018) のように、中間支援組織・技術支援センターが制度化されていたり、十分な公的資金が投入されることが本来は望ましい。

しかし、このような公的な位置づけがなく、予算が十分でなくとも、取り組める範囲で実施可能である点に注目する必要がある。たとえば福祉系大学院など実践家養成系大学大学院が中間支援組織・技術支援センターの役割を担うことはある程度の実現可能性があるだろう。また実践家の職能団体、たとえば日本社会福祉士会や日本作業療法士協会などが、社会的課題の解決のため (たとえば長期入院者の地域移行・地域定着等)、協会をあげて「評価研究チーム」を立ち上げ、実践現場と連携しながら取り組むことも可能であろう。

## 7. おわりに

本稿では、参加型・協働型評価、実践現場を基盤にしたボトムアップ評価の方法を用いて、日常的にプログラムに接する実践家等が積極的・主体的に評価活動に参画・協働して、EBPなどの「効果モデル」を形成・発展させる形成的評価の方法論、PBEE (CD-TEP) 評価法の開発理念、開発プロセスを示し、その結果開発された評価法の概要、基本的な枠組みを整理して提示した。その上で、他の評価アプローチ法との比較からみたこの評価法の意義と特色、有用性、適用範囲と課題について検討し、今後の発展可能性を考察した。

この検討を通して、EBPなど「効果モデル」を形成・発展させる形成的評価法のより良い発展の可能性と、適用範囲から見た形成的評価全般のあり方の検討などの課題が浮き彫りになったと考える。

対人サービスの社会的イノベーションは、実践現場で把握される新しい社会課題への対応と

して誕生して来る。実践現場に関わる実践家等は、社会的イノベーションの「効果モデル」をPBEE (CD-TEP) 評価法など形成的評価の方法論を用いてより効果的なものに形成・発展させ、その科学的な成果を積極的に社会に向けて発信して行くことが期待される。

## 謝 辞

本研究は、JSPS 科研費 JP19203029、JP23243068、JP15H01974の助成を受けたものです。これら研究にご関与頂いたすべての分担研究者、連携研究者、研究協力者、および研究にご参加頂いたすべての実践家、当事者、その他関係者の皆さまのご協力・ご関与で本研究がまとまりました。記して心より深謝申し上げます。

## 参考文献

- 家子直幸 (2022) 「対人サービスにおけるイノベーションをエビデンスに基づく実践プログラムに発展させるための英国の取組み－児童福祉分野のWhat Works Centreを中心に－」、『日本評価研究』、22 (2) : 39-51
- 大島巖 (研究代表) (2022) 「実践家参画型エンパワメント評価の実施基盤を支える有効なEBP 技術支援センターガイドライン構築に向けて」、2022 年度文部科学省・科学研究費補助金基盤研究 (A) 『実践家参画型エンパワメント評価を活用した有効なEBP 技術支援センターモデル構築』 報告書
- 大島巖 (2019) 「社会課題解決に向けた新しい潮流とソーシャルワーク～社会的インパクト評価とEBPMを中心に」、『日本社会福祉学会学会ニュース』、(81) : 14-18
- 大島巖 (研究代表) (2015) 「実践家参画型福祉プログラム評価の方法論および評価教育法の開発とその有効性の検証」、『文部科学省科学研究費補助金基盤研究A研究成果報告書』
- 大島巖、ほか (2012) 「CD-TEP | 円環的対話型評価アプローチ法実施ガイド」、平成22年度文部科学省・科学研究費補助金基盤研究 (A) 『プログラム評価理論・方法論を用いた効果的な福祉実践モデル構築へのアプローチ法開発報告書』 (主任研究者：大島巖)、<http://cd-tep.com/> (2022.4.3取得)
- 大島巖 (研究代表) (2011) 「プログラム評価理論・方法論を用いた効果的な福祉実践モデル構築へのアプローチ法開発」、『文部科学省科学研究費補助金基盤研究A研究成果報告書』
- 大島巖・源由理子 (2020) 「評価の計画：評価デザインと手法の選択」、源由理子・大島巖編 (山谷清志監修) 『プログラム評価ハンドブック～社会課題解決に向けた評価方法の基礎・応用』、晃洋書房
- 大島巖・源由理子・山野則子、ほか (2019) 『実践家参画型エンパワメント評価の理論と方法—CD-TEP法：協働によるEBP効果モデルの構築』、日本評論社
- 木村富美子 (2016) 「社会的企業の支援における中間支援組織の役割」、『創価大学通信教育部論集』、(19) : 15-34
- 津谷喜一郎 (1999) 「EBMとコクラン共同計画」、矢野栄二 (編) 『医療と保健における評価』、南光堂、195-217
- 古矢一郎 (2017) 「政府における「証拠に基づく政策立案 (EBPM)」への取組について」、『季刊行政管理研究』、(160) : 76-85
- 源由理子 (2014) 「地域ガバナンスにおける協働型プログラム評価の試み」、『季刊評価クォーターリー』、7 : 2-17
- 米原あき (2022) 「教育分野における参加型形成的評価の位置付けと意義」、『日本評価研究』、22 (2) : 15-26
- Asmussen, K., Brims L., McBride T. (2019). *10 steps for evaluation success*. Early Intervention Foundation. <https://www.eif.org.uk/resource/10-steps-for-evaluation-success> (2022.4.3取得)
- Blase KA (2009). Technical assistance to promote service and system change. *Roadmap to Effective Intervention Practices #4*. Tampa, Florida : University of South Florida, Technical Assistance Center on Social Emotional Intervention for Young Children.
- Boaz A, et al. (2019). *What works now? : Evidence-informed policy and practice*. Policy Press.
- Chen, HT (2015). *Practical program evaluation*. SAGE.
- Cornell Office for Research on Evaluation (CORE) (2009). *The Evaluation Facilitator's Guide to Systems Evaluation Protocol*. Ithaca, NY : Cornell Digital Services.

- Fixsen DL, Naoom SF, Blase KA, Friendman RM, Wallace F (2005). *Implementation research : A synthesis of the literature*. University of South Florida.
- Gough D, Maiment C, Sharples J (2018). *UK what works centres : Aims, methods and contexts*. EPPI-Centre, Institute of Education.
- IES, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (2011). *National evaluation of the comprehensive technical assistance centers : Final report*. IES.
- Labin SN, Duffy JL, Meyers DC, et al. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, 33, 307-338.
- Littel J.H., White H. (2018). The Campbell Collaboration : Providing better evidence for a better world. *Research on Social Work Practice*, 28 (1), 6-12.
- Patton, M. Q. (2021). Emergent developmental evaluation development. *Journal of Multidisciplinary Evaluation*, 17 (41), 24-34.
- Puttick, R. Ludlow, J. (2013). *Standards of evidence : An approach that balances the need for evidence with innovation*. Nesta.
- [https://media.nesta.org.uk/documents/standards\\_of\\_evidence.pdf](https://media.nesta.org.uk/documents/standards_of_evidence.pdf) (2022.4.3取得)
- Rossi PH, et al. (2004). *Evaluation : A systematic approach (7th edition)*, Sage, 2004 (大島巖他監訳 (2005)『プログラム評価の理論と方法』、日本評論社)
- Senge PM (2006). *The fifth discipline : The art & practice of the learning organization*. Doubleday (=2011, 枝廣淳子ら訳、『学習する組織』、英治出版).
- Wandersman A, et al. (2005). The principles of empowerment evaluation. In Fetterman DM, Wandersman A, eds. *Empowerment evaluation principles in practice*. Guilford Press.
- White H. (2019). *Putting evidence to work, to do what works : the evidence-based project cycle*. Campbell Collaboration.
- <https://www.campbellcollaboration.org/blog/the-evidence-based-project-cycle.html> (2022.4.3取得)
- (2022.6.10 受理)

# **Development of a formative evaluation approach that can foster human service innovations into Evidence-Based Practices Evolvability of the Practitioner and peer Based Empowerment Evaluation using CD-TEP : the PBEE/CD-TEP approach**

Iwao Oshima, PhD

Tohoku Fukushi University  
i-oshima@nifty.com

Kenta Shindo, PhD

Japan College of Social Work  
kentantei@hotmail.com

Yuriko Minamotom, PhD

Meiji University  
yminamoto@mvd.biglobe.ne.jp

## **Abstract**

Various social issues around human services that are not able to take appropriate action have been expanded these days in Japan, and new and effective innovative approaches for the solutions have been explored. For this situation, a new method of formative evaluation that can foster human service innovative “effective models” introduced with new perspectives into more effective Evidence-Based Practices has been needed.

In response to this, authors have developed a formative evaluation named “Practitioner and peer Based Empowerment Evaluation using CD-TEP : the PBEE/CD-TEP approach, which use bottom-up and collaborative evaluation methods based on practices in the fields, and where both practitioners and consumers actively and proactively participate in the evaluation activities.

In this article, we introduced the philosophy and values of the development of the method, and presented the developmental process. In the result, we described outlines and fundamental framework of the PBEE/CD-TEP, and discussed significance, range of application, issues to be solved and evolvability of the method compared with other evaluation methods.

## **Keywords**

formative evaluation, collaborative evaluation, practitioner and peer based empowerment evaluation, bottom-up evaluation, evidence based practice programs (EBPs)

## 【総説】

## 教育分野における参加型形成的評価の 位置付けと意義

米原 あき

東洋大学

yonehara@toyo.jp

### 要 約

教育分野の評価は、歴史的には、学力測定に代表される総括評価としての「測定」から発展し、20世紀前半の「教育測定運動」を経て、より形成的な側面を重視した「評価」へと重点が移行してきた。さらに近年の国際的・国内的な動向に照らしてみると、SDGsの教育目標においても、OECDが提唱する社会情動的スキルの定義やフレームワークにおいても、日本の学習指導要領においても、各種の教育目標がより質的な方向へシフトしているという傾向や、より広い社会的な文脈の中で教育を捉えようとする傾向、そしてより多様な関係者が評価プロセスに参加し協働することが求められているという傾向が共通してみられ、参加型形成的評価に対するニーズが高まっていることが分かる。このような傾向をふまえると、参加型形成的評価の観点から、評価指標やエビデンスについての考え方、そして評価活動がもたらす機能について、再考する必要があると言えよう。

### キーワード

教育評価、参加型形成的評価、測定、協働、評価的コミュニケーション

#### 1. はじめに

教育分野における評価を概説するのは容易ではない。教育分野には重層的な構造、すなわち、学習者—教員—学校—教育行政などの、ミクロ実践からマクロ政策に至る構造があるためである。そのそれぞれのレベルに、学習成果の評価—教員評価—学校評価—教育政策やカリキュラムの評価という大きなテーマが関わっており、加えて、それぞれのレベルに国内外からの様々な議論が影響を及ぼしている。例えば梶田(2010)は、「教育評価の対象となるもの」として14もの対象を挙げている<sup>1</sup>。これらを限られた紙幅の中

で一望しようとするのは野心的に過ぎる試みかも知れないが、一方で、個々のパーツを細かく切り分けて見ることで、教育分野における評価の議論が包括的な視点を失い、テーマ別の方法論に傾倒しがちになっているという点も看過すべきではないだろう。特に2000年以降、教育分野で起こっている革新的な変化を捉えるためにも、巨視的な視点から教育評価を論じる必要があると考えられる。

したがって本稿では、議論の緻密性を一定程度のリスクにさらすことを甘受し、更なる議論の発展に貢献することを期して、教育分野における評価という大きな括りからの総説を試みた

い。同時に、限られた紙幅の中で具体的な議論を展開するためには、ある程度対象を絞る必要がある。本稿においては、国際的に普及が進んでいる初等教育段階に焦点を当てて考察するものとする。

教育分野の評価は、世界史的に見れば、5世紀ごろから始まったとされる中国の科挙制度にまで遡ることができる。日本でも7世紀ごろから大宝令に公務員の登用試験が取り入れられ、その後も社会の発展と共にかたちを変え、対象を拡大しながら現在に至っている(梶田 2010)。教育評価、すなわち、ひとの能力やひとの育成に関わる諸条件の価値を吟味するという営みは、古代より社会の発展を支える重要な役割を果たしてきたと言えるだろう。

古くから人間社会の発展とともにあった教育評価であるが、学校教育制度が整備され、またそれが公教育として普及するにしたがって、教育現場におけるひとつの方法論として見直されるようになった。後述するように、19世紀後半のアメリカにおいて、教育評価は、科学的な研究の対象として位置付けられるようになり、心理学や教育学の発展と共に、20世紀以降、大きな進展を遂げるようになる(梶田 2010; 田中 2010)。本稿では主として20世紀以降の現代における教育評価を考察の対象とする。

本稿の構成は以下の通りである。まず次節(2.)では、アメリカにおける教育評価の変遷に注目し、現代の教育評価が、総括的な「測定」中心のアプローチから、形成的な「評価」へとその重点をシフトしてきた流れを概観する。続いて、そのような流れをうけて、近年注目が集まっている「参加型形成的評価」——すなわち、形成的評価を有効に機能させるために参加型評価の手法を重視する評価のアプローチ——が、どのような位置付けにあるのか、その動向について国際的・国内的な視点から考察する(3.)。最後に、以上の考察を踏まえて3つの論点を抽出する(4.)。総説としての本稿は、これらの論点を検証することではなく、教育分野における参加型形成的評価にかかる試論を提示することを目的とする。

## 2. 教育評価の変遷:「測定」から「評価」へ

### (1) 20世紀初頭の「教育測定運動」

前述の通り、教育分野における評価は、必ずしも量的測定に限らない、「試験」や「考査」による広義の「測定」にその源流がある。20世紀の初頭、科学に対する信頼が急速に高まるなかで、教育においても客観的で恣意性の低い試験の方法が求められるようになった。アメリカ(マサチューセッツ州)では、1845年の教育改革において、それまでは口頭試験で行われていた初等中等教育の評価を筆記試験で行うことになったが、その目的は公平性の担保、すなわち、試験においてすべての児童生徒を同一に扱い、試験官個人の恣意性(えこひいきや歪みの可能性)を低減させることにあった(梶田 2010: 27-28)。

教育評価に客観性を求める動きはその後ますます加速していき、1920年代には「教育測定運動(measurement movement)」となってピークを迎える。この運動を支えたのが、到達度評価や知能テストの発展である。心理学や統計学の発展も手伝って、表1に示すような多様な測定法が開発された。

表1 19世紀末～20世紀初めに開発された教育測定ツール

時期	開発者	測定ツール
1860頃	Fisher	尺度簿(scale-book: 数学や綴字などの教科の評価基準を明示し、5段階の評定に分けたもの) <sup>注1</sup>
1894	Rice	綴字テスト <sup>注2</sup>
1905	Binet	知能テスト <sup>注4</sup> → 1908年日本に導入 <sup>注5</sup>
1908	Stone	算数標準学力試験 <sup>注3</sup>
1910	Thorndike	書き方尺度 <sup>注6</sup>

(注1) Cadenhead & Robinson (1987)、Chadwick (1864)、梶田(2010)、(注2) Engelhart & Thomas (1966)、(注3) Stone (1920)、(注4) Wolf(1973)、(注5) 中村・大川(2003)、(注6) Thorndike (1918)

(出所) 筆者作成

### (2) 1930年代以降の「教育評価運動」

伝統的な教育評価から恣意性や主観性を排除し、公平性や客観性を担保するために取り組まれてきた教育測定運動であったが、様々な批判も生じた。例えばLindquist (1935) は、どんなテ

ストを採用しても、指導の成果の総体を測定することはできないという批判を展開し、かの教育哲学者John Dewey [1859-1952] は、教育測定運動の想定——個人の知能差が教育成果の差の原因であり、教育成果の差が社会階層の差に繋がっているのだとする想定——そのものが「合理的なカースト制度」であると批判した（田中 2010；Gould 1996）。

このような批判を受けて、「測定」の技術的な開発に偏っていた風潮に、「評価（evaluation）」という新たなパラダイムを持ち込んだのが、タイラー原理（The Tyler Rationale）で知られるTyler（1949）である。タイラーは、教育評価の目的は、指導と学習活動の改善にあると考え、「測定から評価へ」の転換を図ったとされる（田中 2010；2013）。タイラー原理とは、以下の4つの問いで構成される評価の考え方である。

- ①学校はどのような教育目的を達成するように努めるべきか。[目標の選択]
- ②どのような教育的経験をを用意すれば、これらの目的が達成できるか。[学習経験の選択]
- ③これらの教育的経験は、どのようにすれば効果的に組織できるか。[学習経験の組織]
- ④これらの目的が達成されているかどうか、どのように判定できるか。[結果の評価]

（田中 2010：29-30）

タイラー原理は、「学校」がどのような教育目的のもとにどのような教育的経験を組織すべきかといった、学校側・教員側の在り方を問うていることが分かる。タイラー原理が想定する評価行為とは、「子どもたちの『個人差』をあばくことではなく、あくまでも授業実践のための『診断的道具』を提供するもの」（田中 2010：26）であり、この点において「教育測定運動」との本質的な違いを読み取ることができる。このような考え方が、「教育測定運動」のオルタナティブとして「教育評価運動（pupil evaluation movement）」を推進する原動力となった（梶田 2010：36）。

### （3）1960年代以降の「形成的評価」

タイラー原理に基づいて様々な研究・実践が展開され、1960年代頃から、その成果がいくつかの教育評価の潮流となってあらわれた。

ひとつは、タイラー原理のような「教育評価」の考え方と、北米や日本の社会でますます影響力を強めていた「教育測定」が融合する動きである。「教育目標の分類体系（taxonomy of educational objectives）」と呼ばれるこのアプローチにおいては、広範に及ぶ教育目標を、認知的領域・情意的領域・精神運動的領域に分類し、各領域において達成すべき目標の系列を明示しようとしたものである（Bloom et al. 1971）。

Tylerの教え子であったBloom（1956；1968）は、当時キング牧師が先導していた公民権運動の支持者でもあり、教育を通じて社会の平等を実現するために、公正かつ科学的で、教育の改善に資するような教育評価の開発に取り組んでいた（田中 2010：49）。Bloomが開発した「分類体系」は、日本を含む世界十数か国で活用されており、現在も教育評価に大きな影響を与えている。

もうひとつの潮流として数えられるのは、形成的な評価の登場である。「教育測定運動」が後押しした学習者のランキング評価によって、評価が学習者の意欲を低減させたり、逆に、当時の優生思想や人種差別といった社会問題と相まって、学習者に不当な優越感を与え得るといった点が問題視されていた。そのような問題点に対する反省として、事後の達成度の測定よりも、評価のフィードバック機能に重点を置いたアプローチ、すなわち、教育活動の改善を目指す形成的な評価のアプローチが提案されるようになった。

例えばCronbach（1967）は適正処遇交互作用（Aptitude Treatment Interaction：ATI）の理論において、教育の成果は、学習者個人の能力のみに因るものではなく、教授方法や教育環境にも依存するという、そしてそれは学習者との交互作用によってもたらされるものであるため、画一的に定式化できない多様性をもつということを論じた。従来、能力測定や知能検査では、学習者個人に対する外的・社会的な要因やそこで生じる交互作用は説明変数として考慮されてこなかったが、ATIの考え方は、「教育実践にありがちな過信、盲信に対する警鐘となる」（並木 1993：119）と言われている。このような視点をもって教育評価をおこなうことで、評価結果の

活用がすすめられ、教授方法や教育環境の改善に向けての議論が一層深められることとなった。

#### (4) 20世紀終盤の「真正の評価」

上述のとおり形成的評価の重要性が叫ばれるようになる一方で、現実の社会では、アメリカにおいても日本においても、学力測定による評価が一層その影響力を強めていた。アメリカでは、1983年に公刊された連邦報告書『危機に立つ国家 (A Nation at Risk)』が教育の危機的状況を訴えると、全国的に教育改革運動が巻き起こった。この運動によって、教育現場には、標準テストの結果による説明責任が強く求められるようになったが、同時に、教育測定に対する批判的な立場からは、標準テスト偏重への懐疑の声が高まった。

このような流れの中で、1980年代後半、アメリカでは「真正の評価 (authentic assessment)」と呼ばれる考え方が登場した。田中 (2013: 26-31) は、「真正の評価」の特徴を以下の6点にまとめている。

- ①評価の対象となる課題や活動が、現実在即したのとなっており、総合力や応用力を必要とするものになっているか。[真正性]
- ②構成主義的な学習観、すなわち、学習者は環境との相互作用の中で意味を構成しながら主体的に学んでいるという学習観を前提としているか。[構成主義的学習観]
- ③学習の結果だけではなく、プロセスも重視しているか。[プロセス重視]
- ④学習成果を評価する方法の開発につとめ、学習者も評価方法を選択することができるか。[方法論の開発]
- ⑤構成主義的な学習観に基づくと、学習における自己調整が最重要視されるが、そのためには学習者の自己評価能力の形成が不可欠となる。学習者の自己評価能力が向上するような評価になっているか。[自己評価能力の形成]
- ⑥真正の評価論では「評価参加者 (stakeholder)」が重要な役割を果たす。評価が、教員や学習者、保護者、地域住民などを含む、参加と共同の実践になっているか<sup>2</sup>。[評価参加者との共同]

「真正の評価」の特徴の中で、従来の教育評価論にはなかった新たな視点として注目すべきなのは、⑥にある「参加と共同」であろう。1980年代の北米の評価学の分野では、ちょうど参加型評価 (participatory evaluation) が興っており (Rossi et al. 2004)、その影響が教育評価の分野にも及んだものと考えられる。従来の教育評価論は、学習者を中心に据えた成果測定の方法論として発展してきており、形成的評価の登場によって教授者 (教員) や教育環境 (学校) へも関心が広げられたものの、地域や保護者といった学外の存在との繋がりまでを広く含めた、「参加と共同の実践」として捉えられてはいなかった。「真正の評価」のこのような特徴は、次節で検討する21世紀の国際的な潮流や日本の動向とも整合する。

### 3. 参加型形成的評価の可能性：国内外の新たな動き

#### (1) 国際的な動向

21世紀の始まりと共に、国連総会でMDGs (Millennium Development Goals) が採択され、その第2目標に「教育」が位置付けられたことは記憶に新しい。MDG2に掲げられた教育目標は「初等教育の完全普及の達成」というもので、これは、1990年の世界教育会議でEFA (Education for All: すべての人に教育を) が採択されて以来、国際社会の共通目標として取り組まれてきた教育の量的拡大を目指すアプローチであった。

MDGsの後継として2015年に採択されたのがSDGs (Sustainable Development Goals) である。SDGsではその第4目標に「教育」が位置付けられ、MDG2の内容が包含されるかたちで、より抽象度の高い目標が目指されることとなった。SDG4の目標は、「すべての人に包摂的かつ公正な質の高い教育を確保し、生涯学習の機会を促進する」と記述されている。

MDG2とSDG4の本質的な違いとして、抽象度の高さと対象範囲の広さが挙げられるだろう (米原 2021)。MDG2が「初等教育」にターゲットを絞り、「完全普及」という量的な拡大を目指すも

のであったのに対し、SDG4は「すべての人」を対象とした、「包括的かつ公正で高質な教育」と「生涯教育」の促進を図ると謳っている。量的拡大を目標とするMDG2の評価は、就学率や識字率といった量的指標を用いることが可能であり、また、これらの指標の達成値が「100%」であることについても暗黙の合意を想定することができた。一方、質的側面を捉えようとするSDG4の評価指標を、普遍的・画一的な量的指標で定義することは非常に困難であり、そもそも評価の対象となるべき「すべての人」とは誰なのかを世界共通の水準で規定することも至難である。結果として、SDG4には11のグローバル指標と32の分野別指標が設定されることとなり（UIS 2021；外務省 2022）、SDG4の目標は各国社会の文脈に応じて「開封（unpack）」する必要があると明言された（UNESCO 2016）。

ユネスコ統計研究所（UIS 2021）が公開しているSDG4の指標リストを見ると、これらの指標には、「比率（proportion）」や「割合（percentage）」として定義されるもののみならず、「（学習成果に関する全国調査の）実施状況 [SDG4.1.6：以下SDG略]」や「（人権教育などの）普及の程度 [4.7.1, 4.7.3]」など、それぞれの社会の文脈の中でしか定義できないような記述的な指標も含まれていることが分かる。

さらに言えば、比率や割合で定義された量的指標でさえも社会的文脈に根差したものが散見される。例えば、「職業技術教育への男女別参加率 [4.3.3]」は、単純に「100%」を目指せばよいというタイプの指標ではなく、どのように達成値を設定すべきか、社会状況に応じた議論が必要となる。あるいは、「地球市民性や持続可能性に関する課題について十分な理解を示している生徒の割合 [4.7.4]」を算出するためには、何を「十分な理解」と定義するのか、何を「持続可能性に関する課題」として教えるのか、といった質的な議論を避けて通れない内容になっている。

このような傾向を示しているのはSDGsにおける教育評価だけではない。経済協力開発機構（OECD）は、2015年に「社会情動的スキル」というキーワードを副題に据えた報告書を公刊し

た（OECD 2015）。社会情動的スキルは、「(a) 一貫した思考・感情・行動のパターンに発現し、(b) フォーマルまたはインフォーマルな学習体験によって発達させることができ、(c) 個人の一生を通じて社会経済的成果に重要な影響を与えるような個人の能力」（OECD 2015：52-54）と定義され、いわゆる認知的なスキルとは異なるカテゴリーとして位置付けられている。例えば、目標の達成に必要な忍耐力や自己抑制、あるいは他者との協働に必要な社交性や思いやりといった能力が社会情動的なスキルと呼ばれる。

認知的なスキルの多くは、OECDの国際学力調査であるPISAをはじめ（OECD 2022）、従来の学力考査などでも測定されてきたが、社会情動的スキルは、認知的スキルよりも測定が困難であると言われている。OECD（2015：56-59）は、「社会情動的スキルを信頼性をもって測定することはできないと仮定する者が多いため、政策論争のなかではこうしたスキルが過小評価されがちである」と警鐘をならし、主観バイアスが生じる恐れを認めつつも、心理学分野で開発された主観指標を活用したり、三角測量などの手法を適用したりすることによって、ある程度の信頼性が確保された測定ができると主張している<sup>3</sup>。

同時に、測定の問題のみにとどまらず、スキル向上のための「学習環境」に対するより包括的なフレームワークも提示されている（OECD 2015：62-64）。このフレームワークによれば、社会情動的スキルは、フォーマルな学校教育だけではなく、ノンフォーマル教育や成人教育も視野に入れた、「家庭・学校・地域社会・職場」というフレームワークで捉えられなければならない。家庭や学校だけではなく、より広範な社会的文脈が——例えば、地域社会のなかでの文化活動や社会的ネットワークの働き、あるいは、職場における職業訓練の機会や労働形態そのものが、「学習環境」として個人の社会情動的スキルの発達に影響を与えるのである。

国連のSDG4やOECDの社会情動的スキルに共通してみられるのは、より広い社会的文脈から教育を捉えようとする姿勢である。学習者を取り巻く環境がますます多様化し、同時に教育を通じて検討すべき社会課題も一層複雑化してい

る現代において、動態的な文脈から学習者個人や特定の能力だけを切り出して評価することには限界があるということだろう。多様な価値、そして刻々と変化する社会的文脈を適切に適時に考慮するためには、多様な関係者の参加と、彼らの間でのコミュニケーションが不可欠となる。UNESCO (2016) が言う「開封」という概念は、参加型の形成的評価によって実体化されると考えられる。

## (2) 国内の動向

では、2000年以降の日本の教育評価にはどのような傾向がみられるだろうか。日本国内の教育、そして教育評価の在り方に最も大きな影響を与えているのは、文部科学省が定める学習指導要領だろう（国立教育政策研究所 2022）。1947年の初版以来、およそ10年ごとに改正されている指導要領だが、2000年に入ってから、2003年、2007年、2017年に改正・告示が行われている<sup>4</sup>。

2003年の改正では、従来の学力に偏重した学校教育に対する批判を受けて、「生きる力」の育成が教育目標として掲げられ、学校完全週5日制が始まった。いわゆる「ゆとり教育」の導入期である。続く2008年の改定では、当時のゆとり教育批判に応じるかのように、「ゆとりでも詰め込みでもなく、『バランスのとれた生きる力』」への転換が謳われている。同時に、上述のMDGsや、2005年から開始された国連ESD10年などの影響を受けて<sup>5</sup>、「持続可能性」というワードが指導要領に初めて登場したり、「国際（理解）」や「多様（性）」といったワードの出現頻度が急増するなどの新たな傾向も見られた（米原 2014）。さらに2015年の指導要領一部改訂では道徳の教科化が図られるなど、社会情動的なスキルの育成が目指されていることも読み取れる。

2017年に告示が出され、2020年ごろから順次施行がすすめられている、いわゆる新指導要領もこのような流れの延長上に位置づく（文部科学省 2017）。認知的学力だけではなく、非認知的能力を伸長するためには、「主体的・対話的で深い学び」が必要であるとされ、アクティブラーニング（AL）が推奨されるようになった。ALは

教科横断的・単元横断的な学習として取り入れられることが多いため、効果的なALを行うためには、カリキュラム・マネジメントと評価が鍵となる。『小学校指導要領解説：総則編』第3章第3節「2 学習評価の充実」（文部科学省 2017：93-94）には、児童参加型の評価を推奨する記述がみられる（下線部筆者）。

教師による評価とともに、児童による学習活動としての相互評価や自己評価などを工夫することも大切である。相互評価や自己評価は、児童自身の学習意欲の向上にもつながることから重視する必要がある。（中略）このような資質・能力のバランスのとれた学習評価を行っていくためには、指導と評価の一体化を図る中で、論述やレポートの作成、発表、グループでの話し合い、作品の制作などといった多様な活動を評価の対象とし、ペーパーテストの結果にとどまらない、多面的・多角的な評価を行っていくことが必要である。

この記述には、「意欲の向上」のために参加型の評価が必要であること、そのためにはプロセス（指導）と評価が一体化する必要があること、そして、認知テストのみに限らない多様な評価の手法を導入する必要があることが明示されている<sup>6</sup>。また、同第3章第1節「4 カリキュラム・マネジメントの充実」（文部科学省 2017：40）には、学校全体のマネジメントや、地域や家庭といった学校外のステイクホルダーとの関連についても述べられている（下線部筆者）。

各学校においては、各種調査結果やデータ等に基づき、児童の姿や学校及び地域の現状を定期的に把握したり、保護者や地域住民の意向等を的確に把握した上で、学校の教育目標など教育課程の編成の基本となる事項を定めていくことが求められる。

ここでは、エビデンスに基づき、学校外の関係者とも連携しながら、教育課程の編成（カリキュラム・マネジメント）をすすめていくことが推奨されている。

このような新たなカリキュラム・マネジメントを実現するための体制として、「社会に開かれた教育課程」という考え方のもと、コミュニティ・スクールの構想が具体的に示されている（文部科学省 2021）。コミュニティ・スクール構想においては、学校運営に地域の声を反映するための「学校運営協議会」や、学校と地域が協働して多様な教育活動を実施するための「地域学校協働本部」などが組織され、「地域と学校が、子供たちの学びの充実のために、協議し、協働し、活動後の評価をして、また次の取組につなげていく」というPDCAサイクルを回していくことが重要」（文部科学省 2021）と謳われている。

日本社会に学校評価が本格導入されたのは2000年ごろであったが、当初は反論も多く、また、学校評価が学校に「評価疲れ」をもたらしているという指摘も少なくなかった。現在においてもそのような傾向が払拭されたとは言い難いところはあるが、コミュニティ・スクール構想などに具体化される参加型の形成的評価によって、多様な学校関係者をつなぎ、信頼関係を構築し、学校をエンパワーしようという声が、近年高まっている。しかしながら、エンパワーメント評価や参加型評価の成功事例が着実に増加している一方で（天笠・大脇 2012；池田・池田 2018；橋本・岩永 2020；米原 2021など）、アンケート調査によるデータ収集や学校運営協議会の場が形骸化している例も少なからずあり、日本の学校現場では、本質的な評価が主流化しているとは言い難い。

#### 4. 3つの論点

教育評価の歴史的な変遷と、2000年以降の国内外の動向を概観すると、次のような特徴を指摘することができるだろう。

まず、学習者に求められる能力の本質が変わってきている。20世紀に想定されていた教育評価の対象が主として認知的なスキルであったのに対し、OECDのいう社会情動的スキルや日本で教科化された道徳のように、より非認知的な能力への関心が高まっている。また、SDGsなどに

みられるマクロな議論においても、量的な目標からより質的な目標への重点の移行が認められる。

次に、より広い社会的な文脈の中で教育活動を捉えようとする傾向が強くなっていることが指摘できる。SDGsの指標群、OECDが提示するフレームワーク、そして近年日本で推進されているコミュニティ・スクール構想などにその傾向が見て取れる。より広い文脈に関心を払うことにより、必然的に関係者の拡大と多様化が起こっており、それに付随して、多様な関係者間での、価値をめぐる本質的なコミュニケーションの重要性が高まっている。

これらの特徴が、教育評価の議論に何をもたらすのだろうか。ここでは参加型形成的評価の観点から、3つの論点について考察したい。

##### 論点① 評価指標の考え方

非認知的な能力や質的な目標への関心が高まることにより、従来の定量指標による評価が困難になっているというのは既述の通りである。どのような能力、あるいはどのような社会状況が「望ましい」状態なのかを一元的に定義できる場合は、その社会における普遍性の高い価値を指標にあらわし、その指標を用いて総括評価を行うことで、公正な評価が実現されるだろう。例えば、「教育測定運動」では、より公正な評価のために統一的な学力考査による測定が行われた。また、国際的な教育開発目標においては、より普遍的な評価を行うために就学率などの世界共通の指標が採用された。

これに対して、非認知的スキルや質的目標は、一元的な定義や統一的な試験で測定することが困難である。SDG4の指標群は40を超え、社会情動的スキルの評価においては主観指標の活用も奨励されている。形成的な評価の重要性が見直される傾向にある一方で、形成的な評価ほど評価指標を設定することが難しいというジレンマがここにあらわれている。

ひとつの試論として、特に形成的評価においては、指標も形成的に捉えるというアプローチがあっても良いのではないだろうか。総括的評価における評価指標の意義は、達成度の可視化

と成果の比較可能性にあると考えられるが、形成的評価における指標の意義として、変容過程の可視化と評価的なコミュニケーションの活性化を想定できるとすれば、関係者の間で独自の指標を「手作り」することの意義も生まれるのではないだろうか。つまり、評価指標を現場の関係者たちが形成する過程で目標や価値をめぐる議論が交わされ、その指標を参加型評価の中で活用することによって自らの変容過程をモニタリングすることができ、その結果として、評価活動に対する関係者のオーナーシップが高まり、自分たちにとってより納得感の深い、現場の改善に役立つ評価活動が行われる可能性がある（米原 2019）。

### 論点② エビデンスの考え方

上記の論点と関連して、エビデンスの捉え方についても柔軟なアプローチがあり得るのではないかと考えられる。医療分野などで言われる科学的エビデンスは、「バイアスのない方法により得られたデータを、バイアスのない方法で分析して得られた結果の総体」（正木・津谷 2006：12）と定義されるが、現象学や哲学においては、個人が自らの体験や経験を省察した際に「確かにこうである」と確認できる確実性や不可疑性のことをエビデンス（明証性）と定義し、科学的エビデンスと区別して「反省的エビデンス」と呼んでいる（小林・西 2015）。また今井（2015）は、生活者の日常経験や生活感覚と必ずしも合致しない「近代科学的なエビデンス」と、学習者の生活経験に根拠づけられた「生活世界的なエビデンス」を区別し、教育分野における「学習の経験」は、生活世界的なエビデンスでしか把握できないと指摘する。

ここで試論的に、社会科学における科学性・客観性の要求水準を、自然科学と人文科学の間地点に想定できるとすれば、主観を排したデータに基づく科学的エビデンスと、個人の主観的内省や生活経験を根拠として採用する反省的・生活世界的エビデンスの間地点に、「合意に基づくエビデンス」という概念を展開できないだろうか。すなわち、論点①に述べたような、ある取り組みの関係者の間で合意された——客観

でも主観でもない間主観的な——質的・量的指標によって収集されたデータを、社会的なエビデンスと考える立場があっても良いのではないだろうか。

### 論点③ 評価による協働と協創

SDG4、OECDのフレームワーク、そして日本のコミュニティ・スクール構想に共通して見られたのが、社会的な文脈への関心と、それに伴う関係者の多様化であった。関係者の多様化とはすなわち価値の多様化を意味し、そこでのコミュニケーションの難しさと重要性が同時に浮かび上がってくる。UNESCO（2016）が推奨するような「開封」を実現するために、参加型形成的評価が協働と協創の「場」としての機能を担い得るのではないかと考えられる。

参加型形成的評価は、日常的にはコミュニケーションの機会を持つことがない、異なる立場の関係者を繋ぐプラットフォームになり得る。また、そこには評価活動という目的があり、ある具体的な取り組みについて議論を交わすという舞台装置が整っている。例えば、コミュニティ・スクール構想の中の学校運営協議会などは、異なる立場の関係者が、評価を通じて、価値をめぐる本質的なコミュニケーションをとる場としての機能を持っている。このような場には、その学校や地域の文脈に相応しい指標を検討したり、そこから得られたエビデンスをもとに改善案を検討したりするといった潜在的な可能性がある。つまり、学校運営協議会は意見を具申するための会議体としてだけでなく、共に新たな価値を創出する協働と協創（co-production）の場として機能し得る。

現状ではこのように機能している事例は多くないかも知れないが、参加型形成的評価の考え方が教育評価の一環として広く認識されることによって、学校運営協議会のような潜在性をもつ現場が、評価的思考とそれに基づく評価的なコミュニケーションの場として、その潜在能力を発揮できるようになるのではないだろうか。

## 5. おわりに

本稿では、教育評価の歴史的な変遷を概観することによって、現代の教育評価の特徴を描出し、3つの論点を提示した。これらの論点をふまえ、理論・方法論上の課題と、実践上の課題を指摘して本稿の結びとしたい。

まず理論的な課題について、本稿では、教育分野における参加型形成的評価への関心が高まっていることを指摘し、参加型形成的評価がもつ潜在的な可能性について言及した。しかしながら、文脈への依存性が高く、画一的にマニュアル化し難い評価のアプローチと、SDGsや全国的な政策などのマクロなアプローチとをどのように繋ぎ、整合させていくのかについては、十分な先行研究が蓄積されておらず、総説として検討することができなかった。

私見としては、メゾレベルの公共機関（学校教育の例でいえば教育委員会などの組織）における参加型評価のコミュニケーション機能に期待したい。すなわち、ミクロな教育現場からマクロな政策現場に至る関係者が集える場をメゾレベルで提供し、参加型評価の手法を活用して三者の間で評価的なコミュニケーションを醸成するということである。ここで重要なのは、評価的なコミュニケーションを形骸化させないことと、間主観的な「合意に基づくエビデンス」を活用することである。行政の会議にありがちな上意下達の情報伝達の場とならないようなファシリテーションの工夫が必要である。この点については、以下の実践上の課題とも併せて、今後更なる事例の蓄積と理論化が必要であろう。

また、上記の理論化と並行する方法論上の課題として、①より質的な指標や記述的な指標についての研究開発、②形成的な指標策定の方法論に関する研究蓄積、そして③マクロな政策指標とミクロな形成的指標の接合に関する研究の必要性についても付言したい。

次に実践上の課題について、本稿では、近年の教育評価においては、多様な関係者の参加が求められる傾向が強くなっていることを指摘し、参加型形成的評価が、立場の異なる関係者を繋ぐコミュニケーションの「場」としての機能を

持ちうることを述べた。しかしながら、評価に関わる関係者が増加するという事は、それだけ多くの人々が評価的思考を持つ必要があるということでもある。教育の現場で言えば、特に教員が評価的思考を持っていることが不可欠となるだろう。今日の教員は、ALなどの多様な教育活動に合わせて多様な学習評価の手法を使いこなせる必要があることに加えて、学校運営協議会のような評価的コミュニケーションの場ではファシリテーターを任されることが多い。その場の潜在的な可能性をどれだけ引き出せるかは、ファシリテーション・スキルの良し悪しに左右されるということになる。

しかしながら、現状の教員養成課程では、評価学の概論や評価的思考について修得する機会が設定されていない。学習成果の評価方法についても多様化が進んでいるが（西岡 2011；松下 2012）、それらの方法論に加えて、評価的思考やファシリテーション技術などを含めた、広義の評価学と評価実践を学修する機会を提供する必要があるだろう。先進的な教育委員会では、これらに類する研修を提供している例もみられるが、今後、教員養成課程の見直しから検討する必要があるかも知れない。

以上のとおり、試論に基づく理論・方法論上の課題と実践上の課題を指摘したが、本稿の冒頭に述べたように、教育評価は古くから人間社会の発展と共にあり、社会の変化に応じて、また時には社会の変化を先導したり後押ししたりしながら、その在り方を変容させてきた。今日の教育評価は、予測不可能な時代の教育をどのように捉え、改善していくのかという問いに挑戦することを期待されているのではないだろうか。

## 注記

1 教育評価の対象として以下の14項目が挙げられている。①学習者個々人、②教育活動、③教育内容（カリキュラム）、④教師（教育指導者）、⑤学習者集団、⑥教師をも含めた学級（ホームルーム）、⑦教師集団、⑧学校全体のあり方、⑨基本的施設、⑩校地および校舎、⑪地域的環境、⑫教育施設の管

- 理・運営、⑬教育行政システム、⑭社会全体における各種教育施設の位置付けと機能・役割（梶田2010：6-11）。
- 2 「共同」には「協働」「協同」などの表記もあるが、ここは田中（2013）に依拠している。本稿では出典や筆者の意図によってこれらの表記を使い分けている。
  - 3 各国の取り組み事例として、カナダで開発された「社会的責任」の能力基準や、ベルギーで開発された「ウェルビーイング」の測定指標、ノルウェーで実施されている「社会的・情動的な満足度」評価のための調査などがある（OECD 2015：148-152）。
  - 4 改正・告示の年次は校種によって若干のずれがある。本文の年次は義務教育段階を参照した。また、施行年次は改正・告示の年次と異なる。詳細は国立教育政策研究所（2022）の学習指導要領データベースにて確認されたい。
  - 5 2005-2014年にかけて実施された国連の『持続可能な開発のための教育（ESD）10年キャンペーン』を指す。小泉内閣時代に日本が提案し、国連総会にて全会一致で採択された国際教育キャンペーンである。
  - 6 ペーパーテストに限らない多様な学習評価の手法も、2000年以降、開発が進んでおり、ポートフォリオ評価（西岡2011）やパフォーマンス評価（松下2012）がその代表例である。

## 参考文献

- 天笠茂・大脇康弘（2012）『学校をエンパワーメントする評価』、ぎょうせい
- 池田琴恵・池田満（2018）「エンパワーメント評価型学校評価の導入における校長の意識の変容過程」、『教育心理学研究』、66（2）：162-180
- 今井康雄（2015）「教育にとってエビデンスとは何か：エビデンス批判をこえて」、『教育学研究』、82（2）：188-201
- 外務省（2022）『SDGグローバル指標：4 質の高い教育をみんなに』、  
（<https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/statistics/goal4.html>、2022年3月25日閲覧）
- 梶田毅一（2010）『教育評価』（第2版補訂2版）、有斐閣双書
- 経済協力開発機構（OECD）（2019）『社会情動的スキル：学びに向かう力』、明石書店
- 国立教育政策研究所（2022）『教育研究情報データベース：学習指導要領の一覧』、（<https://erid.nier.go.jp/guideline.html>、2022年3月25日閲覧）
- 小林隆児・西研（2015）『人間科学におけるエヴィデンスとは何か：現象学と実践をつなぐ』、新曜社
- 田中耕治（2010）『教育評価』、岩波書店
- 田中耕治（2013）『新しい「評価のあり方」を拓く：「目標に準拠した評価」のこれまでとこれから』、日本標準
- 中村淳子・大川一郎（2003）「田中ビネー知能検査開発の歴史」、『立命館人間科学研究』、6：93-111
- 並木博（1993）「教授・学習研究におけるATIパラダイムと適性理論」、『教育心理学年報』、32：117-127
- 西岡加名恵（2011）『教科と総合に活かすポートフォリオ評価法：新たな評価基準の創出に向けて』、図書文化
- 橋本洋治・岩永定（2020）「学校教育への保護者参加の効果と参加意欲に関する研究」、『現代と文化』、140：21-31
- 正木朋也・津谷喜一郎（2006）「エビデンスに基づく医療（EBM）の系譜と方向性：保健医療評価に果たすコクラン共同計画の役割と未来」、『日本評価研究』、6（1）：3-20
- 松下佳代（2012）『パフォーマンス評価：子どもの思考と表現を評価する』、日本標準
- 文部科学省（2017）『小学校学習指導要領（平成29年告示）解説：総則編』、文部科学省
- 文部科学省（2021）『社会に開かれた教育課程の実現に向けて』、（<https://manabi-mirai.mext.go.jp/torikumi/chiiki-gakko/syakaini-hirakareta.html>、2022年3月25日閲覧）
- 米原あき（2014）『「国際理解教育」から『持続可能な開発のための教育』へ：学習指導要領にみる変遷と展開』、『ガバナンス研究』、10：197-216
- 米原あき（2019）「協働型プログラム評価実践における『協働型社会調査』：参加型のアプローチでエビデンスを創出する試み」、『評価クォーターリー』、50：2-17
- 米原あき（2021）「SDG教育ゴールにみる理念型目標の評価に関する一考察：測定可能性（measurability）から評価可能性（evaluability）へ」、『日本評価研

- 究』、21 (2) : 31-46
- Bloom, B. S. (1956) . *Taxonomy of educational objectives: The classification of educational goals* [Cognitive domain - Handbook 1]. Mackay.
- Bloom, B. S. (1968) . Learning for mastery. *The Evaluation Comment*, 1 (2) . (=稲葉宏雄・大西匡哉監訳 (1986) 『すべての子どもにたしかな学力を』、明治図書 [所収])
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971) . *Handbook on formative and summative evaluation of student learning*. McGraw-Hill. (=渋谷憲一ほか訳 (1989) 『学習評価ハンドブック』、第一法規出版)
- Cadenhead, K., & Robinson, R. (1987) . Fisher's "Scale-Book" : An early attempt at educational measurement. *Educational Measurement: Issues and Practice*, 6 (4) , 15-18.
- Chadwick, E. (1864) . Statistics of educational results. *Museum: A Quarterly Magazine of Education, Literature, and Science*, 3, 479-484.
- Cronbach, L. J. (1967) . How can instruction be adapted to individual differences? In R. M. Gagne (Ed.) , *Learning and individual differences*. Merrill.
- Engelhart, M. D., & Thomas, M. (1966) . Rice as the inventor of the comparative test. *Journal of Educational Measurement*, 3 (2) , 141-145.
- Gould, S. J. (1996) . *The mismeasure of man*. W. W. Norton & Company.
- Lindquist, E. P. (1935) . Cooperative achievement testing. *The Journal of Educational Research*, 28 (7) , 511-519.
- OECD. (2022) . *Programme for International Student Assessment [PISA]*. Retrieved March 25, 2022, from <https://www.oecd.org/pisa/>
- OECD. (2015) . *Skills for social progress: The power of social emotional skills*. Retrieved March 25, 2022, from <https://www.oecd.org/education/skills-for-social-progress-9789264226159-en.htm>
- Rossi, P., Lipsey, M., & Freeman, H. (Eds.) . (2004) . *Evaluation: A systematic approach* (7th ed.) . Sage publications.
- Stone, C. W. (1920) . No fairy tale: The need for standardized tests. *Journal of Educational Research*, 2, 868-869.
- Thorndike, E. L. (1918) . *Seventeenth yearbook of the national society for the study of education: Part II*. Public School Publishing Company.
- Tyler, R. W. (1949) . *Basic principles of curriculum and instruction*. University of Chicago Press. (=金子孫市監訳 (1978) 『現代カリキュラム研究の基礎』、日本教育経営協会)
- UNESCO. (2016) . *Unpacking sustainable development goal 4 education 2030 guide*. UNESCO.
- UNESCO Institute for Statistics [UIS] (2021) . *SDG 4 global and thematic indicator lists: Official list of SDG4 indicators*. Retrieved March 25, 2022, from <https://tcg.uis.unesco.org/sdg-4-global-and-thematic-indicator-lists/>
- Wolf, T. H. (1973) . *Alfred Binet*. University of Chicago Press.

(2022.6.6受理)

## **Trend and Dynamics of Participatory Formative Evaluation in the Field of Education**

Aki Yonehara

Toyo University  
yonehara@toyo.jp

### **Abstract**

The history of educational evaluation is rooted in “measurement” for a summative evaluation such as achievement examination. However, after the educational measurement movement in the early twenties, the emphasis was shifted to more formative aspect of evaluation. Reflecting on the current global and domestic trend, some significant changes can be observed in common among the SDGs' educational evaluation, the OECD's definition and framework of Social Emotional Skills, and the National Course of Study in Japan. These changes include the tendencies: first, the purpose of education has changed into more qualitative direction; second, the concept of education has been captured in the wider social context; and lastly, more diverse stakeholders need to participate and coproduce in the evaluation process. Such new tendencies indicate the increasing needs for participatory formative evaluation and require us to reconsider the meanings of evaluation indicator and evidence and the function of evaluation activities.

### **Keywords**

Educational evaluation, Participatory formative evaluation, Measurement,  
Coproduction, Evaluative communication

## 【総説】

## 社会的インパクト評価の系譜 — マネジメント支援のための評価への進化 —

今田 克司

(一財) CSOネットワーク

katsuji@csonj.org

### 要 約

近年、「社会的インパクト評価」について論じられることが増えてきた。一方で、評価研究の中で、この概念が定着しているとは言い難い。本論では、社会的インパクト評価の概念整理を試み、国内外における理論や実践の議論が、概念を成長・進化させていると論じる。具体的に取り上げるのが、2010年代中盤以降の北米（特に米国）におけるインパクト投資とインパクト測定・マネジメント（IMM）の動きと、日本国内における内閣府社会的インパクト評価ワーキンググループによる概念整理である。これらを概観し、新たに見出された社会的インパクト評価の特質として、学び・改善、マネジメント支援や意思決定のための評価の側面が強調されていることを述べる。

### キーワード

社会的インパクト評価、インパクト測定、インパクト投資

#### 1. はじめに一混同と混乱

「社会的インパクト評価」。日本国内で2010年代中盤ぐらいから話題にのぼるようになった用語である。これが普及するようになった大きなきっかけは、内閣府が2015年度に当時の共助社会づくり懇談会のもとに社会的インパクト評価ワーキンググループを設置し、2016年3月に報告書「社会的インパクト評価の推進に向けて—社会的課題解決に向けた社会的インパクト評価の基本的概念と今後の対応策について」（内閣府2016）を発表したことといえよう（付記）。ワーキンググループの報告書と同時に、社会的インパクト評価に関する調査研究報告書も内閣府の委託調査結果として発表されている（三菱UFJリ

サーチ&コンサルティング株式会社2016）。

この調査研究報告書を始めとして、社会的インパクト評価に関して参照された海外の文献のうち、概念整理や実践枠組みの提供で大いに参考にされたものに、英国の非営利団体、New Philanthropy Capital (NPC)の資料がある（Benedict, Lumley and Ní Ógáin 2011; Ní Ógáin, Lumley and Pritchard 2012; Kazimirski and Pritchard 2014）。当時、NPCでは、英国の非営利セクター（チャリティ）が自分たちの成果をインパクト測定（impact measurement）によって適切に提示することを支援する活動を行っていた。これらNPCの資料では、impact measurement と evaluation の用語が使われているが、おもな主張はインパクト測定による活動の成果の開示で、これを日本では「社

会的インパクト評価」として紹介していた。

加えて、海外の投資活動から2000年代後半に生まれたインパクト投資において、インパクト測定は、投資先の企業が展開する事業や投資行動自体の成果の報告・開示においても取り沙汰され始めていた。グローバルに展開する投資行動の規範に変化が生じ、ESG投資やサステナブル・ファイナンスが脚光を浴びる中で、投資における財務的リターンとリスクに加えて、社会的環境的インパクトを適切に計測しようとする機運が高まり、これが日本では「社会的インパクト評価」または単に「インパクト評価」と呼ばれるようになっていった。

日本・海外含め、評価研究においては言わずもがなのことであるが、インパクト測定による成果の開示は、評価研究が対象とする評価＝エバリュエーション (evaluation) 活動の一断面に過ぎない。また、インパクト評価 (impact evaluation) は、因果命題である介入 (原因) がアウトカム (結果) に対して及ぼす影響、すなわちインパクトを推定することである (津富・正木 2020: 106)。津富・正木 (2020) が指摘するように、「社会的インパクト評価」では、因果関係の推定は必ずしも求められておらず (2020: 111)、上記のような経緯で「社会的インパクト評価」なるものが日本において取り沙汰されることとなったことは、評価研究との関係でいえば混同と混乱の種まきから始まったこととなる。

## 2. 2つの世界の乖離と接合

### (1) 「インパクト測定」と「評価研究・実践」の世界

上記のように、日本語においては、翻訳の問題もあり、社会的インパクト評価 (メジャーメント) がインパクト評価 (エバリュエーション) と混同されるという事態が生まれたが、米国においても、公的セクターや国際開発、財団のフィランソピー等で定式化されてきたエバリュエーションと、民間セクター (特に、投資・金融セクター) に起源をもつインパクト測定 (メジャーメント) がいかに接合するのかについて、

2010年代中盤ぐらいから議論が活発になっていった。

そもそも、「インパクト測定」すなわちメジャーメントは、社会課題解決における民間セクターの役割が注目されたことにより、介入の効果を市場メカニズムの中でいかに「評価」するかという発想から生まれている。投資・金融関係者の実務的なニーズに応じるもので、市場の中で評価を一般化し、効果的な取り組みが流通していくような仕組みを充実させていくことを主眼にしている。そこでは、比較できることが大きな要素であり、標準化やレーティングの試みが必要不可欠となる (今田 2018a)。

一方、エバリュエーションはより学術的で、その主たる目的は説明責任を果たす (特に社会的介入による成果が仮説通り生まれたか)、学び・改善を図る (評価によって事業体自身が学び、成長していくか)、新たな知識を創造する (評価から一般化できる要素を抽出し、共有知として活用できるか) のいずれかあるいは複数にある。評価専門家コミュニティを形成し、研究、実践の両面から分野の発展を担っている。

評価専門家コミュニティから見れば、メジャーメントの世界にエバリュエーションの知見を注入することで、社会課題解決への大きな貢献が見込まれるために、2つの世界の接合に力を入れる動きが起こった。すなわち、伊藤・玉村・植野 (2021)、塚本・関 (2020)、今田 (2020) などが紹介するように、インパクト測定は社会的投資の投資効果を測る手段として開発されてきたが、米国の評価研究者・実践家にとっては、これは評価研究の知見を十分に活用することなく進化してきたものなのである。Reismanら (Reisman et al. 2015) は、2015年に発表した小冊子において、インパクト測定には、おもに投資家に情報を提供するために社会・環境インパクトに関する指標を開発し、それによって投資効果を的確に測定しようという意図があると解説している。そちらの世界から見ると、評価実践は、

- ・「事後」の評価であるために事業改善に役に立たない
- ・投資判断に役立てる即応性がない

- ・ 難しい言葉で書かれた分厚い報告書である場合が多く、簡便に利用できない
- ・ 変化の激しい現実社会に適合していない固定的モデルで行われる

と見え、そこで自分たちで社会的インパクトの評価モデルを作ろうという動きにつながった<sup>2</sup>。

翌年の2016年、アメリカ評価学会 (AEA: American Evaluation Association) の年次大会と前後するように、「インパクト・コンバージェンス (インパクトの収束)」なる会議が開催された。評価関係者にインパクト測定に関心をもっても

らおうとした試みで、年次大会で新たに社会的インパクト測定に関するワーキンググループが学会内に設置されている (通称、SIM-TIG)<sup>3</sup>。しかし、学会全体から見ると、評価研究の知見をインパクト測定という新しい分野に活用してもらうために積極的に働きかける目論見は、必ずしもうまくいかなかったという<sup>4</sup>。また、インパクト投資コミュニティなど、インパクト測定を推進する側からも、評価専門家やその知見を積極的に活用しようとする動きはすぐには活発にならなかったという<sup>5</sup>。

表1 評価とソーシャル・インパクト測定の対比

評価	ソーシャル・インパクト測定
公的セクター、国際協力・開発、財団のフィランソロピーに起源をもつ。	民間セクター (特に、投資・金融セクター) に起源をもつ。
測定・評価の目的には多種多様なものがあり、それは評価ユーザが評価を何に利用する意図をもっているかによって決まる。大きな括りとしては、(1) 説明責任の確保、(2) 学びや改善、(3) 知識の創造があげられる。	測定の目的はリターンを決定し、意思決定に活用するため (経済的リターンが主だが、昨今は環境、社会などのリターンも考慮するようになってきている)。
複雑な事象からいかにデータを集め、分析し、理解し、価値判断を下すか、その一連の作業のすべてに価値がある (評価プロセスは評価結果と同様に重要)	意思決定者 (多くの場合は投資家) がスピード感をもって決定できるように、複雑な事象をできるだけ単純化して捉えることが求められる。測定結果に価値がある。
データや情報を公開・共有し、それを様々なアクターが活用できるようにして、共有財産とすることに価値を認める。	データや情報には市場価値があり、その売買も成立するので、データや情報を共有するには抵抗感が伴う。

(出所) Reisman, et al. (2015) より筆者が編集。今田 (2018a) に掲載。

## (2) 評価とインパクト測定の橋渡しの試み

このような流れの中で、評価関係者にメジャーメント側の動きに関心をもってもらい、エバリュエーション側の知見を提供する機運を盛り上げようという意図もあり、AEAのSIM-TIGメンバーが中心になって執筆し、2018年、AEAの学会誌である American Journal of Evaluation (AJE) に、「インパクト測定と評価の接合点 (Where Impact Measurement Meets Evaluation)」をテーマに特集号が生まれ、5本の論文が掲載された。

特集号の冒頭の小論で、Vo and Christie (2018) は、2016年の「インパクト・コンバージェンス」会議に言及した上で、この接合点について再考する意義について記述している。「2つの世界には共通のルーツもある」(2018:385) とした上で<sup>6</sup>、「最近までは、ほとんどお互いに独立して存在していた」(2018:385) と、2つの世界の距離感をあらためて指摘している。

Ruff and Olsen (2018) は、「インパクト投資家は、場所、組織、活動分野が何であれ通用する、一連の統一された測定基準を求める」(2018:402) と指摘し、これは「誤った探求である」(2018:402) と述べている。インパクト測定にとって、異なる介入間でインパクトの質量を比較できるような標準化を進め、市場が投資判断をできるようにすることは必須課題だが、評価専門家の立場から言えば、「ちょっと待った」なのだ。それは、リングとオレンジを比べるようなもので、標準化へと安易に進めてはならないことは、介入の文脈を大切に評価研究の蓄積が教えている。インパクト測定の分野が抱えるこのようなニーズに対し、いかに評価者コミュニティが回答を差し出すことができるのか、評価者がこれを課題として引き受けなければならないという。

同様に、Harji and Jackson (2018) は、評価研究

者や実践家が、民間投資や社会的投資活動という比較的新しい分野で自分たちの力を発揮するためには、「民間セクターのアクターの言語を理解することが不可欠である」(2018: 399)と説く。社会課題解決のアクターが多様化し、民間企業が参入しているという現代的背景においては、公的セクターや民間助成財団による助成事業の評価に加え、民間セクターによる社会課題解決の分野で評価者が活躍の道を探るのは必要不可欠だという。

### (3) 評価の知見が提供する価値—IMMの定着

さきに、評価とインパクト測定の世界の乖離が、2016年以降すぐには進まなかったという識者の見解を紹介した。しかし、AJEの特集号が発行された2018年からさらに時を経て、接合は進みつつある。ここではIMMの定着に注目してみたい。

インパクト測定におけるここ数年の進展でもっとも目覚ましいものとして、IMM (インパクト測定・マネジメント: Impact Measurement and Management) をあげることができる。インパクト投資におけるIMMの進化と定着において評価研究者・実践家が果たした役割の大きさについては確固たる定説があるわけではない。しかし、前述のAEAのSIM-TIG関係者含め、評価専門家が有益なインプットをしてきたのがこの分野であることは間違いない。

例えば、SIM-TIGの有力メンバーであり、オックスフォード大学サイドビジネススクールで、インパクト・メジャメント・プログラムの運営をしているカリム・ハルジ氏は、HOW (いかにインパクトを測定するのか) の前にWHY (なぜ測定するのか) とWHAT (具体的に何を測定するのか) を問わなければならないと説いている (今田 2021:108-109)。

インパクト投資のエコシステム・ビルダーとして知られるGIIN (Global Impact Investing Network) が2020年に出したIMMに関する調査報告書 (第2版) によれば、第1版が出てから3年の間に、「IMMの実践や報告についての投資家の理解」、「IMMについてのガイダンス」、「IMMに関するツールやフレームワークの洗練」、「IMMの

実践に関するリサーチやデータ」、「IMMのスキルをもつ専門家」において、「大いに」または「ある程度」進んだと考える回答者がほぼ9割または9割以上という結果が出ている (Global Impact Investing Network 2020:26)。

GIINは、IMMを「ビジネス上の行動や投資が人や地球に与えるプラスとマイナスの両方の影響を特定し検討すること、そして、マイナスの影響を緩和し、プラスの影響を最大化する方法を、自身の目的に照らし合わせながら考え出すプロセス」と定義している<sup>7</sup>。IMMとは、インパクト創出のための一連のプロセスであり、インパクト測定はそれに内包されている。さきに見た、Vo and Christine の整理では、インパクト測定の関心対象となるユニットはおもに「活動 (activities)」と「結果 (output)」だったが、IMMの定着により、それは評価の関心対象である「投入 (input)」、「活動 (activities)」、「結果 (output)」、「成果 (outcome)」へと広げられたと理解できる (Vo and Christine 2018:386, 伊藤・玉村・植野 2021:100)。

IMMのやり方を定式化したものとしては、国際金融公社 (IFC) からスピノフした Impact Principles が作成した「インパクト投資運用原則」<sup>8</sup> と国連開発計画 (UNDP) が作成した「SDGインパクト基準」<sup>9</sup> がここ1-2年で知られるようになってきている。後者は、事業者に対するインパクト認証を行う計画もあり、日本の企業関係者の関心も高い。2020年12月に行われたオンライン説明会には300人を超える参加があったが、この説明会で、SDGインパクト基準 (スタンダード) が、「実践スタンダードであり、パフォーマンス・スタンダードや報告スタンダードではない」と注意喚起がされた (今田 2021:110)。その意味するところは、問われるのが一定のインパクトを創出しているかではなく、インパクトを創出するためのマネジメントができていくかという点だということだ。標準化しようとしているのは、IMMの結果ではなく、IMMのやり方そのものである。

GIINによるIMMの定義のすぐあとも、「IMMは本来、反復的なものである」<sup>10</sup>という解説が見られる。この「反復的 (iterative)」という形容詞

は、インパクト投資関係者との意見交換などで IMM の特徴としてよく出てくる単語になっている。それは何よりも、IMM が、単発的なものでも直線的（リニア）なものでもないことを表している。

### 3. 日本国内における社会的インパクト評価の展開－内閣府の社会的インパクト評価ワーキンググループ

以上のように、海外特に北米においては、2010 年代中盤より、評価とインパクト測定の橋渡しの試みが始められ、それは今も続いている。ここ数年で、2つの世界の接合はある程度進んだと言えるが、それはまだ一部の動きという見方もある。

ここで日本国内に目を転じて、社会的インパクト評価の発端となった内閣府のワーキンググループにおける整理を概観してみよう。このワーキンググループについては冒頭に言及したが、伊藤・玉村・植野（2021）は、「社会的インパクト評価として定義される手法の多くは…プログラム評価の多様な側面を包含するもの」（2021: 91）、「社会的インパクト評価も、事業の実施プロセスや成果を明らかにし、社会状況を改善するための活動に資する情報を提供するという意味で、プログラムの評価の一類型であると理解することができる」（2021: 94）と述べている。

実際、筆者も参加したこのワーキンググループにおいて、社会的インパクト評価とプログラム評価の関係性は議論になった<sup>11</sup>。特に第2回会合の議事要旨を読み返すと、社会的インパクト評価がメジャメントかエバリュエーションかという議論があり、結論的に「メジャメントやエバリュエーションという二者択一というよりは、両方必要であると思っている」という発言が残っている<sup>12</sup>。

最終的な報告書（内閣府 2016）において、「社会的インパクト」および「社会的インパクト評価」は次のように定義されている。

社会的インパクト: 短期、長期の変化を含め、当該事業や活動の結果として生じた社会的、

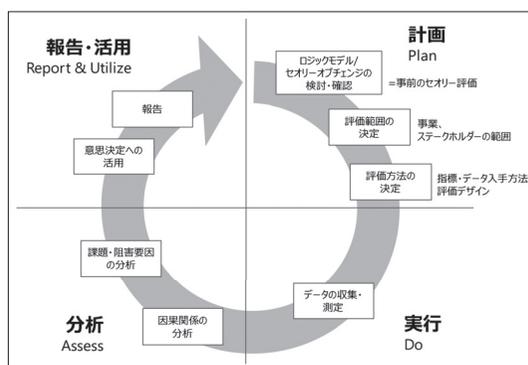
環境的なアウトカム

社会的インパクト評価: 社会的インパクトを定量的・定性的に把握し、当該事業や活動について価値判断を加えること（内閣府 2016: 2）

また、「評価の目的、活用、効果・意義」の説明では、「[評価]は[監査]、[査定]ではなく、「価値を引き出す」こと」（内閣府 2016: 2）との記述が見られ、「価値（value）を引き出す（接頭辞 e(x)-）」ことを原義にもつエバリュエーションを意識したものであることがうかがえる。

さらに、評価の過程（プロセス）を、1計画、2実行、3分析、4報告・活用の4つのステップから構成されると、図1の図示とともに説明されていることを見ても、すでにこの段階でプログラム評価の色合いが濃くなっていることがわかる。

図1 社会的インパクト評価のプロセス



（出所）内閣府（2016: 14）

### 4. 新たに見出された社会的インパクト評価の特質

以上のように、2010年代中盤以降、社会的インパクト評価をめぐっては国内外で大きな動きが見られている。北米においては、IMMの定着において評価専門家が活発なインプットを行い、インパクト測定にエバリュエーションとマネジメントの視点と実践を統合しようとしている。日本国内においては内閣府のワーキンググループを発端に、プログラム評価の考えや要素を意識的に取り入れた展開が見られる。例として、休眠預金活用事業における「資金分配団体・実行団体に向けての評価指針」（日本民間公益活動

連携機構 2020) や社会的インパクト・マネジメント・イニシアチブ (SIMI) の「社会的インパクト・マネジメント・ガイドライン」(社会的インパクト・マネジメント・イニシアチブ 2021)、GSG国内諮問委員会による「インパクト投資におけるインパクト測定・マネジメント実践ガイドブック」(GSG国内諮問委員会IMMワーキンググループ 2021) の特徴などについて、別稿で論じることとしたい。

共通しているのは、社会的インパクト評価を成果やインパクトの可視化という「点」の測定作業にとどめず、マネジメントという一連の「線」の活動における連続した作業と位置付けたことで、よりよい意思決定や改善のために評価を行うという考え方が定着しつつあることだろう。

では、ここまでの議論をもとに、新たに見出されている社会的インパクト評価の特質を指摘してみたい<sup>13</sup>。

### (1) 改善のツールとしての評価

まず、評価研究において、事業終了後に主に説明責任を果たす目的で行われる評価を総括的評価、それ以前の段階で事業改善に役立てるために行われる評価を形成的評価とタイプ分けすることが多い(源 2020:26)。特に日本国内の社会的インパクト評価において、プログラム評価の考え方を導入して4つのステップを踏んでいく評価プロセスが内閣府ワーキンググループの報告書で示されたことは、形成的評価で志向される学び・改善を社会的インパクト評価の中核に据えたことを物語っている。

### (2) マネジメント支援としての評価

事業改善を目途とした形成的評価と重なる部分も多いが、IMMや社会的インパクト・マネジメントの流れに見られるように、国内外で、社会的インパクト評価は明らかにマネジメントや組織の意思決定支援へと大きくシフトしている。「実用重視」の評価という評価目的を打ち立てたのはPattonだが(Patton 1997)、「役に立つ評価」という視点から評価をマネジメントに直結させようという動きが加速している。これは、社会的インパクト評価に直接関係する動き以外の場

面でも2010年代中盤から見て取れ、それは、総括的、形成的評価に加え評価目的の「第三の柱」と呼べるような動きになっていた(今田 2014)。その後、社会的インパクト評価においてこの動きが吸収されつつあり、意思決定支援のための評価という評価の役割が注目されている。

### (3) リアルタイムの評価

次の点はマネジメント支援としての評価の一部と整理してもよいだろうが、意思決定支援のためには評価情報が意思決定に間に合うタイミング、すなわちリアルタイムで提示されなければならない。情報技術の革新で、これは可能になってきている。簡単な例では、アンケートをオンラインでやることにより、自由回答などのテキスト情報も含め、結果が即時に集計され、意思決定者すなわち事業実施者にフィードバックとして返ってくることはもはや珍しいことではない。

そのような状況において求められる情報は、「こういう事実がわかっていたらあの時こういう意思決定をしていたのに」と後からわかるような情報のことである。インパクト・マネジメントの実践によって、それが後追いではなく、「現在進行形」で活用できるかもしれないというのがインパクト・マネジメントと社会的インパクト評価の一つの「売り」になる(今田 2018b)。Pattonはこれについて、「あなたが今知らないことで、もし知っていたら、あなたが今行っていることと違った決断や異なる行動をとる情報は何か？」という言い方をしている(今田 2019)。Pattonは、事業実施者の決断の瞬間を分岐点(a fork in the road)と呼び、事業実施者は状況把握に努め、課題解決のための有効な打ち手を考えるが、その場面場面に伴走し、事実やその分析を提示するのが評価者の役割だという(今田 2019)<sup>14</sup>。

### (4) フロントエンドの評価

次の点もマネジメント支援としての評価の一部と整理してもよいだろうが、インパクト投資においては、評価を役に立たせようとする場面はフロントエンド、特に投資の意思決定におい

て現れる。組成・ストラクチャリング段階において一般的なIMMステップがもっとも細かく刻まれているのはそのため、投資家は投資先候補である事業者と、インパクト情報をいつどうやって入手、報告するか、そのための一連のプロセスはどのようにするか、投資家と事業者がどのようなエンゲージメントによって事業の進捗をモニタリングするかを決めていく。

### (5) セオリーの重視と柔軟な運用

これまでの特質からもわかるように、フロントエンドでは、的確なインパクトテーマ、インパクトゴールを設定できるかが重要な通過点となる。そもそもインパクト投資は投資家や事業者の「意図」を前面に押し出すものだが、そうであれば、事業によってどの分野の何の目標を達成したいのかを明確に描き、それをセオリーで提示することが求められる。アウトカムや指標やインパクトKPIの設定、データ収集・分析は、すべて設定されたセオリーに紐づくので、まずはセオリーの部分がしっかりできていないと適切な評価もマネジメントも望むことはできない。

そして、そのセオリーは、組織内外の状況の変化に応じて柔軟に運用されるものでなければならない。その際、リアルタイムのフィードバックが役に立つ。社会的インパクト評価やインパクト・マネジメントは、仮置きでもよいのでセオリーを可視化し、それを関係者で共有することの重要性を説いている。それによって共通のゴールに向けた事業運営が可能になるからだ。

加えて、IMMにおいては、ステークホルダーとのエンゲージメントが強調されている。投資家が「意図」をもち、それをもとにインパクト・ゴールを定め、投資を実施するという一連の流れの中で、どのようなステークホルダーとどこにエンゲージメントするかによって、事業の質に大きな影響が出るという考え方が流通している<sup>15</sup>。

## 5. おわりに－混乱から発展へ

これまで、社会的インパクト評価は、「学術的

な研究よりも、社会的事業の実践において発展してきたという経緯もあり、必ずしも評価研究の中で中心的な議題として取り上げられてこなかった」(伊藤・玉村・植野 2021:90)。今日においても、社会的インパクト評価は確立された評価アプローチではないし、まだしばらくは過渡期・成長期としての動きが続くだろう。しかし、当初の混乱を経て、それは極めて現代的な社会課題解決の要請に答えるものとして発展しつつある。

本論では社会的インパクト評価について十分な考察をなし得たわけではないが、この現在進行形の試みに、評価研究に携わる専門家や実践者が多く理論と実践の両面から参入することが望まれている。今後さらに議論や整理が必要なテーマとしては、次が考えられる。

- ・リアルタイムの評価、フロントエンドの評価、セオリーの重視も含め、社会的インパクト評価の総括的评价と形成的評価のバランスや力点の置き方はいかに考えればよいのか。そもそも、総括的评价と形成的評価の二分法が、社会的インパクト評価の理解の上で適切なものなのか。
- ・インパクト投資における事業者のインパクト・マネジメントにインパクト投資家が伴走する場面や、社会的インパクト・マネジメントの評価支援者が事業者に伴走するなど、評価者がリアルタイムで伴走する構図が描かれるようになってきている。そこでは、そもそも伴走型の評価アプローチである発展的评价が有効と考えられるが、発展的评价に関しては、海外に比べて日本国内の概念整理や実践例づくりが進んでいない。これをいかに効果的に進めることができるのか。
- ・評価に関わるアクターが多様化し、社会的インパクト評価においても多くの事業関係者の参加が構想されていることから、評価専門家そしてそれ以外の関係者にとって必要とされる評価的思考はいかなるもので、それはいかに高めていくことができるのか<sup>16</sup>。また、これと関連して、関係者の「評価のキャパシティビルディング (ECB)」にはどのようなアプローチで臨めばよいのか。

このような問いを含め、社会的インパクト評価がより多くの研究者や実践家によって掘り下げられていくことによって、現代社会で複雑化・深刻化している社会課題が解決の方向に向かう後押しとして貢献していくことが望まれている。

## 謝辞

本論執筆にあたり、大澤望氏（株式会社大沢会計アンド人事コンサルタンツ）、加えて査読者の方々には有益なご指摘をいただきました。この場を借りて、感謝いたします。

## 付記

参考文献等で明らかにした以外にも、筆者は本論で紹介した2015年度内閣府共助社会づくり懇談会社会的インパクト評価ワーキンググループの委員として関わっている。透明性の確保のためにも、それを記しておきたい。

## 注記

- 1 <https://www.npo-homepage.go.jp/kaigi/kyoujo-shakai/kyoujo-shakaitoha#2803> (3/25/22 アクセス)
- 2 Reisman et al. (2015) をもとに、筆者による2016年日本評価学会全国大会発表資料「社会的インパクト評価における評価伴走者の役割～発展型評価の知見から」を加工した。
- 3 Social Impact Measurement TIG (Topical Interest Group)。日本からはソーシャル・バリュー・ジャパンの伊藤健氏と筆者が参加した。
- 4 当時のAEA会長で、社会的インパクト・マネジメント・イニシアチブ (SIMI、当時は社会的インパクト評価イニシアチブ) が2018年の年次イベント、Social Impact Day で基調講演者として招いた John Gargani 氏のコメント。今田 (2018a) 参照。
- 5 Jane Reisman氏含め、SIM-TIGメンバーとのディスカッションより。ただし、これは2018年ぐらいまでで、その後、後述のように、インパクト投資関係者から評価研究を積極的に活用する動きが見られる。
- 6 共通点や相違点が表にまとめられており、伊藤・玉村・植野 (2021:100) により、日本語での紹介がさ

れている。

- 7 <https://thegiin.org/imm/> (筆者による仮訳)
- 8 <https://www.impactprinciples.org/sites/default/files/2021-06/Impact%20Principles%20Brochure%20Revised.pdf>, [https://www.ifc.org/wps/wcm/connect/fe499630-792d-434f-8dd2-f5d06da4c1ed/Impact+Investing+Principles\\_+FINAL.pdf?MOD=AJPERES&CVID=mSUxyEd](https://www.ifc.org/wps/wcm/connect/fe499630-792d-434f-8dd2-f5d06da4c1ed/Impact+Investing+Principles_+FINAL.pdf?MOD=AJPERES&CVID=mSUxyEd) (参考和訳) (いずれも3/25/22アクセス)
- 9 <https://sdgimpact.undp.org/practice-standards.html>, [https://www.jp.undp.org/content/tokyo/ja/home/library/sdgimpact\\_standard\\_enterprises.html](https://www.jp.undp.org/content/tokyo/ja/home/library/sdgimpact_standard_enterprises.html) (企業・事業体版の日本語訳) (いずれも3/25/22アクセス)
- 10 <https://thegiin.org/imm/>
- 11 なお、本論においてプログラム評価とは、「社会課題解決をめざす社会的目的を達成するための社会的介入プログラム (social intervention program) のプログラム」(源 2020:20) を評価対象とするものと位置付ける。
- 12 <https://www.npo-homepage.go.jp/uploads/h27wg-impact-hyouka-2-gijiyoushi.pdf> p.17
- 13 「社会的インパクト評価」には、そもそも英語からの翻訳の問題があった。本論で整理した新しい特質をもつ「社会的インパクト評価」をどのように英語に翻訳するかも議論のポイントとなるだろう。SIMIでは、英文サイトなどで、これを Evaluation for Social Impact と訳している。
- 14 これは特に発展的評価について言えることである。発展的評価については Patton (2011) など参照。
- 15 例えば、SIMIも参加した OECDのGlobal Action: Promoting Social and Solidarity Economy Ecosystems におけるワーキンググループ、Putting Stakeholders in the Center 参照。  
<https://socialvalueus.net/sse-plp> (3/25/22 アクセス)
- 16 評価的思考に関しては、今田 (2020:220-221) などを参照。

## 参考文献

- 伊藤健・玉村雅俊・植野準太 (2021) 「プログラム評価の一類型としての「社会的インパクト評価」の課題と可能性」、『日本評価研究』、21 (2):89-101
- 今田克司 (2014) 『海外の非営利組織事業評価の動向に関する情報収集』、田中弥生科学研究費調査基盤

- (B) 「公共の境界領域の評価～政府、企業と非営利組織の輪郭」の一環として行われた海外の非営利組織の事業評価に関わる動向調査  
(<https://www.csonj.org/images/de1b9a75d34735a7ee058a07cd5bfa56.pdf> 3/25/22アクセス)
- 今田克司 (2018a) 「「取り扱い注意！」のソーシャル・インパクト」、オルタナ・オンライン  
(<http://www.alterna.co.jp/24843> 3/25/22アクセス)
- 今田克司 (2018b) 「インパクト・マネジメントは活用へのシフト」、オルタナ・オンライン  
(<http://www.alterna.co.jp/25088> 3/25/22アクセス)
- 今田克司 (2019) 「ソーシャル・イノベーション支援としてのDE」、ブルーマーブル・ジャパン 『DE引き出し集』 ([https://www.blue-marble.co.jp/docs/a02/b02/#\\_ftn3](https://www.blue-marble.co.jp/docs/a02/b02/#_ftn3) 3/25/22アクセス)
- 今田克司 (2020) 「NPO事業評価」、山谷清志・他編 『プログラム評価ハンドブック』、晃洋書房、212-224
- 今田克司 (2021) 「寄稿: インパクト投資とIMM」、GSG国内諮問委員会 『日本におけるインパクト投資の現状と課題: 2020年度調査』、107-112  
(<http://impactinvestment.jp/user/media/resources-pdf/gsg-2020.pdf> 3/25/22アクセス)
- GSG国内諮問委員会IMMワーキンググループ (2021) 『インパクト投資におけるインパクト測定・マネジメント実践ガイドブック』 ([https://impactinvestment.jp/user/media/resources-pdf/Guidebook\\_for\\_Impact\\_Measurement\\_and\\_Management.pdf](https://impactinvestment.jp/user/media/resources-pdf/Guidebook_for_Impact_Measurement_and_Management.pdf) 3/25/22 アクセス)
- 社会的インパクト・マネジメント・イニシアチブ (2021) 『社会的インパクト・マネジメント・ガイドライン Version 2』 (<https://simi.or.jp/wp-content/uploads/2021/07/Social-Impact-Management-Guideline-ver.2.pdf> 3/25/22 アクセス)
- 塚本一郎・関正雄 (2020) 『インパクト評価と社会イノベーション—SDGs時代における社会的事業の成果をどう可視化するか』、第一法規
- 津富宏・正木朋也 (2020) 「プログラムの介入後の評価—インパクト評価の方法—」、山谷清志・他編 『プログラム評価ハンドブック』、晃洋書房、106-133
- 内閣府 (2016) 『社会的インパクト評価の推進に向けて—社会的課題解決に向けた社会的インパクト評価の基本的概念と今後の対応策について』  
(<https://www.npo-homepage.go.jp/uploads/social-impact-hyouka-houkoku.pdf> 3/25/22 アクセス)
- 日本民間公益活動連携機構 (2019、2020改訂) 『資金分配団体・実行団体に向けての評価指針』  
([https://www.janpia.or.jp/hyouka/download/hyouka\\_sisin\\_2020.7.pdf](https://www.janpia.or.jp/hyouka/download/hyouka_sisin_2020.7.pdf) 3/25/22 アクセス)
- 三菱UFJリサーチ&コンサルティング株式会社 (2016) 『社会的インパクト評価に関する調査研究最終報告書』 (<https://www.npo-homepage.go.jp/uploads/social-impact-hyouka-chousa-all.pdf> 3/25/22 アクセス)
- 源由理子 (2020) 「「プログラム評価」とは」、山谷清志・他編 『プログラム評価ハンドブック』、晃洋書房、19-30
- Global Impact Investing Network (2020). *The State of Impact Measurement and Management Practice, The Second Edition*. ([https://thegiin.org/assets/GIIN\\_State%20of%20Impact%20Measurement%20and%20Management%20Practice\\_Second%20Edition.pdf](https://thegiin.org/assets/GIIN_State%20of%20Impact%20Measurement%20and%20Management%20Practice_Second%20Edition.pdf) 3/25/22 アクセス)
- Harji, Karim and Edward T. Jackson (2018). Facing Challenges, Building the Field: Improving the Measurement of the Social Impact of Market-Based Approaches, *American Journal of Evaluation*, 39 (3), 396-401.
- Kazimirski, Anne and David Pritchard (2014). *Building your Measurement Framework: NPC's Four Pillar Approach*. New Philanthropy Capital. (<https://www.thinknpc.org/wp-content/uploads/2018/07/NPCs-four-pillar-approach-FINAL.pdf> 3/25/22 アクセス)
- Ní Ógáin, Eibhlín, Tris Lumley and David Pritchard (2012). *MAKING AN IMPACT: Impact measurement among charities and social enterprises in the UK*. New Philanthropy Capital.  
(<https://www.thinknpc.org/wp-content/uploads/2018/07/Making-an-impact.pdf> 3/25/22 アクセス)
- Patton, Michael Quinn (1997). *Utilization-Focused Evaluation: The New Century Text, 3rd Edition*, Thousand Oaks, Sage Publications. (大森彌監修、山本泰・長尾眞文編訳 (2001) 『実用重視の事業評価入門』、清水弘文堂書房)
- Patton, Michael Quinn (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, New York, NY: Guilford Press.

- Reisman, Jane, et al. (2015). *Streams of Social Impact Work: Building Bridges in a New Evaluation Era with Market-Oriented Players at the Table*. The Rockefeller Foundation Evaluation Office. (<http://orsimpact.com/directory/streams2rev.htm> 3/25/22 アクセス)
- Reisman, Jane, Veronica Olazabal, and Shawna Hoffman (2018). Putting the “Impact” in Impact Investing: The Rising Demand for Data and Evidence of Social Outcomes, *American Journal of Evaluation*, 39 (3), 389-395.
- Rickey, Benedict, Tris Lumley and Eibhlín Ni Ógáin (2011). *A JOURNEY TO GREATER IMPACT: Six charities that learned to measure better*. New Philanthropy Capital. (<https://www.thinknpc.org/wp-content/uploads/2018/07/A-journey-to-greater-impact.pdf> 3/25/22 アクセス)
- Ruff, Kate and Sara Olsen (2018). The Need for Analysts in Social Impact Measurement: How Evaluators Can Help. *American Journal of Evaluation*, 39 (3), 402-407.
- Vo, Anne T. and Christina A. Christie (2018). Where Impact Measurement Meets Evaluation: Tensions, Challenges, and Opportunities. *American Journal of Evaluation*, 39 (3), 383-388.

(2022.6.27 受理)

## **From Social Impact Measurement to Evaluation for Social Impact: Evolution of the Concept and Practice towards Evaluation for Management Support**

Katsuji Imata

CSO Network Japan

katsuji@csonj.org

### **Abstract**

Since around 2015, “social impact measurement” has become a topic of discussion among many in Japan who are working to address contemporary social issues, especially by those in the private sector. At the same time, the concept is far from being firmly established in the evaluation discipline. This paper attempts to organize the concept of social impact measurement and argues that discussions of its theory and practice in Japan and abroad have led to the growth and evolution of the concept. Specifically, the paper will focus on how it is evolving since the mid-2010s by paying attention to the field of impact investing and impact measurement and management (IMM) in North America (especially in the U.S.), as well as how it has been treated in Japan by examining the work of the Working Group on Social Impact Measurement of the Cabinet Office. The result is a conceptual re-framing of its measurement focus to a learning, improvement and management focus. Lastly, the paper discusses the prospects for this approach to evolve into a new form of management and decision-making support with an evaluative framing.

### **Keywords**

evaluation for social impact, impact measurement, impact investing



## 【実践・調査報告】

# 対人サービスにおけるイノベーションをエビデンスに基づく 実践プログラムに発展させるための英国の取組み ー児童福祉分野の What Works Centre を中心にー

家子 直幸

三菱UFJリサーチ&コンサルティング

ieko@murc.jp

## 要 約

英国では中央政府が2000年前後から社会政策でエビデンスを重視する方針を推進し、対人サービスについてもエビデンスに基づく実践が要請されてきた。児童福祉分野でのエビデンスに基づく取組みは対人援助職の専門性向上を図る改革と相まって進められ、中でも早期支援・早期介入といった予防的政策への期待は強い。緊縮財政により政策の焦点が大規模実証からエビデンスの活用へと移る中、主として有効性に関するエビデンスを政策や実践の意思決定で用いるよう支援する知識仲介機関、What Works Centreが中央政府主導で設置されてきた。What Works Centreの基本的機能はエビデンスの創出・翻訳・適用であり、児童福祉分野では対人サービスでのイノベーションをエビデンスが検証された実践プログラムへと発展させられるよう、形成的評価の各段階と資金提供とを関連付けたアプローチが取られている。

## キーワード

対人サービス、エビデンスに基づく政策・実践、予防的政策、知識仲介機関、形成的評価

### 1. はじめに

多岐にわたる公共政策において、住民への現物給付として直接的に提供される人的サービス（対人サービス）は、福祉行政の中で大きな割合を占める。本稿が着目する児童福祉分野でも、福祉・心理・保育・教育・保健・医療・司法といった多様な専門的背景を有する職種が、ソーシャルワーカーやケアワーカー等の対人援助職として対人サービスの提供を担い、クライアントとなる子どもやその保護者等へ個別性に応じた支援を行っている。

児童福祉分野の政策や実践の特徴として、子

どもや家庭が抱える問題は複合的で複雑性が高いこと（ムンロー 2021）に加えて、クライアントの対人サービスの利用が非自発的なケースも比較的多いこと、政策や実践<sup>1</sup>の裏付けとして研究以外にも児童の権利に関する条約等が理念的根拠とされること、子どもの発達には主要なアウトカム領域が学際的で広範囲にまたがること（Early Intervention Foundation 2019）等が挙げられる。また、対人援助職の臨床的知識・経験の重要性が保健医療（特に臨床サービス）等よりも強調される分野であるとされ（ナトリー 2015）、エビデンスに基づく実践（evidence-based practice）は他分野に比して進んでいない（Boaz

et al. 2019)<sup>2</sup>。

本稿では、このような児童福祉分野における対人サービスで、クライアントのアウトカムを一貫して向上させるとのエビデンスが検証された実践プログラム、EBPプログラム（大島ら 2019）を取り入れようとする継続的な取組みが見られる英国（特にイングランド及びウェールズ）を取り上げる。英国では児童福祉分野の中でも予防的政策に重点を置いた展開を図っており、対人サービスにおけるイノベーションを「(改善が既存の資源をより効率的に用いることでより良い結果を得るのに対し) 新たな資源あるいは既存の資源の新たな利用によってより良く異なる結果を得る、主流の方法論を一変させるような新規性のある実践やモデル」(Department for Education 2017, p.16) と定義して、イノベーションを段階的にEBPプログラムへと発展させるイニシアティブを進めてきた蓄積がある。特に近年、英国内でエビデンスに基づく政策や実践を牽引しているWhat Works Centreと呼ばれる機関の取組みに着目することで、対人サービスにおけるイノベーションと形成的評価の位置付けや役割を検討する。

## 2. 英国中央政府の児童福祉分野における動向と予防的政策のイニシアティブ

ここでは、英国の中央政府が2000年前後から政府全体としてエビデンスを重視する方針を強く打ち出し、対人サービスに関してもエビデンスに基づく実践への政策的期待感が高まりをみせたこと、その前後で官僚主義に陥った対人援助職の専門性向上を図るための児童福祉改革が目指されたこと、予防的政策を推し進める中央政府のイニシアティブが変容しつつも現在まで引き続き行われていることを、それぞれの観点から概観したい。

### (1) 政策と実践におけるエビデンスの重視

エビデンスに基づく政策が英国内で本格的に導入されたのは、トニー・ブレア労働党政権が成立した1997年以降だとされる（家子ら 2016）。

ブレア政権では、グローバル化した市場において国際競争力を確保するためにも、子どもの貧困が将来のライフチャンスを損ねることのないよう、また社会的排除や反社会的行為、そして世代間の連鎖につながらないように、人的資本への投資が重要だと位置付けられた（Eisenstadt and Oppenheim 2019）。1999年に剥奪・貧困地域での地域プログラムとして60箇所で開催され、翌2000年には約500箇所、2010年までにはシュア・スタート子どもセンター（Sure Start Children's Centre）として約3,500箇所へと拡大されたシュア・スタート事業は、労働党政権によるエビデンスに基づく社会的投資を具現化した看板政策と位置付けられて推進された。

中央政府の政策決定においてエビデンスを重視する潮流は、2010年に保守党を中心とした連立政権への交代後も現在に至るまで継続している。2008年のリーマンショック以降は財政再建が至上命題となり、今日まで緊縮財政政策が取られているが、これは対人サービスにおいても例外ではない。また同時に、児童福祉に限らず教育や医療の分野でも中央政府から地方自治体、そしてさらに委託等を通じて非営利組織や社会的企業へと権限移譲が進められ、クライアントと直接接する現場により近いところで意思決定されるように行政改革が進められた（小池 2020）。その際、財政的制約がある中で効率的に対人サービスが提供されるよう、何が有効か（what works）に関してその時点で利用可能で最良のエビデンスに焦点が当てられるようになった。これは、エビデンスを活用する局面、つまり政策の立案・決定段階だけでなく、実践においてもエビデンスが強く意識されるようになったことを意味する。

### (2) エビデンスに基づく実践と児童福祉改革

対人サービスの実務家にエビデンスに基づく実践を求める一方、中央政府内では新公共経営（NPM：New Public Management）の手法を展開しており、業績測定によるコントロールが対人サービスの現場にも大きな影響を及ぼした。この点に関しては「合理的／客観的／科学的なアプローチがソーシャルワーク専門職に浸透してき

た時期は、ベストバリュー（マネジリアルイズム）が徹底され出した時期と重なる。すなわち、合理的／客観的／科学的なアプローチは、行政管理／経営を強化することで、最高のアウトカムの達成を目標むベストバリューの脈絡に呑み込まれ、資源的脈絡や個別援助の質への配慮はなおざりにされた」（田邊 2019、p.340）との指摘にも表されるように、子どものために良かれと考えて導入したはずの成果志向の枠組みがかえって現場の官僚化を招いたとされる。

そのような折、児童福祉分野では2000年に続いて2007年にもロンドン北部の同じ地区で凄惨な児童虐待死亡事件が起きたことで、対人援助職に対する社会的圧力が極めて厳しくなり、現場が委縮した結果、クライアントとの距離が遠ざかってしまった。この状況に警鐘を鳴らし、対人援助職の専門性が尊重されるシステムを再構築するように児童福祉改革の必要性を訴えたのが、2011年の政府報告（Department for Education 2011、通称「ムンロー・レビュー」）である。ムンロー・レビューでは15の勧告が示されたが、そこでは過剰なコンプライアンスやマニュアル作業といった行動様式を脱却し、実践を担っている対人援助職自身が知識向上や批判的省察によって継続的に学習し、エビデンスに基づいた判断や実践を自律的に行うことで支援の質の改善を図る姿が提唱された（ムンロー 2021）。このビジョンは、現在の英国の児童福祉分野の各種政策へと受け継がれている。

### （3）早期支援イニシアティブとその変遷

既述のとおり、英国では実践の局面におけるエビデンスの重要性を再認識しつつも、それがNPMと相まって意図せずNPMの強化に寄与した結果、対人サービスの質の低下を招いた。その状況を乗り越える方策として、エビデンスを踏まえた省察的実践の重要性が示され、今日までの児童福祉改革の理念的基礎となっている。ムンロー・レビューの勧告では他にも、より国内外のエビデンスに着目すべきことや、エビデンスを踏まえた早期支援が重要であることにも言及している。

早期支援のイニシアティブは、こと労働党政

権下ではコミュニティを基盤とする家庭支援を、プログラム化した予防的政策として展開することで対人サービスにイノベーションを起こすべく、中央政府から地方自治体へ多額の資金が投入された（Eisenstadt and Oppenheim 2019）<sup>3</sup>。しかしながら、そのような早期支援のイニシアティブは主として米国で創出されたエビデンスに着想を得て開始されたものの、プログラムの設計段階や実装段階で中核的要素に変更が加えられた。例えば、シュア・スタートは当初、学際的な知見に基づいた理論的背景を踏まえてプログラムが設計された。

- ・ 貧困・学業不振・社会的排除の世代間連鎖が子どものライフチャンスに影響を与える
- ・ 幼少期の子どもの成長は、これまで考えられていた以上に環境の悪影響を受けやすい
- ・ 乳幼児とその家族に提供される行政サービスの質のバラつきが大きく、かつ統合・調整されていない

そのため剥奪・貧困地域の4歳未満の子ども・親（英国内の子どもの20%）が対象だったが、厳密な評価が行われないうまま政治的要求から対象地域が拡大され、かつエビデンスに忠実な実装方法では展開されず各地域の資源や優先順位に応じた多様なアプローチが許容された（Early Intervention Foundation 2020a）<sup>4</sup>。英国では「予防は治療に勝る」との一般的合意があり、早期の集中的なアプローチへの投資が子どもや家族により良い結果をもたらし、長期的にはコスト削減にもつながると考えられているものの、「予防的政策」という表現には曖昧さが伴うため修辭的に用いられやすい等、不確実性が高い（Cairney and Denny 2020）。シュア・スタートをはじめ早期支援を推進するイニシアティブにおいて、このように中央政府からの独立性が保たれずに実装が進められたことで、エビデンスを踏まえた早期支援を掲げながらも実態は政策に基づくエビデンス（policy-based evidence）だとする批判もなされた（Boaz et al.2019）。

2010年以降の保守党政権では、公共政策への資金拠出は大幅に削減され、対人サービスでエビデンスを活用する動きにも変化がみられた。イノベーションにつながりうるアイデアを大

規模に展開し、実証研究によってエビデンスを直接創出するアプローチから、先行文献レビューや二次研究といったより低廉な研究手法で、成功している諸外国の取組みから学んで技術移転するアプローチへと徐々に焦点が移行していく(Boaz et al.2019)。例えば、ムンロー・レビューを契機として創設されたChildren's Social Care Innovation Programmeは、2014-2016年の第1期で57プロジェクトに計1億1,000万ポンド、2016-2019年の第2期では50プロジェクトに対し2億ポンドが提供され、サービスの質向上、費用対効果の改善、イノベーションの要因特定を目的に実施された(Department for Education 2017、2020)。評価はこのプロジェクトの中核をなすと位置付けられ、形成的評価と総括的評価が別立てで行われている。このうち形成的評価では、実施主体の要請に応じて外部の評価チームが適宜フィードバックを行ったり、対人援助職が実践の学習過程を振り返るワークショップを開催したりするなど、評価キャパシティ形成(大島ら 2019)を通じた研究の知識と臨床の知識を統合するための試みが行われるようになっていった。

### 3. 児童福祉分野のWhat Works Centre

前節でみてきたように、2000年前後以降の労働党政権と保守党政権の双方でNPMへの取組みがみられる中、労働党政権はエビデンスを社会的投資へ活用する観点から予防的政策を推進したのに対し、緊縮財政政策をとる保守党政権は早期支援イニシアティブを継承しつつ、大規模実証プロジェクトの実施よりも対人サービスの実践の局面におけるエビデンス活用の重要性を強調した。以降では、主に保守党政権以降で強化された中央政府の取組みとして、対人サービス等の公共政策で有効性に関するエビデンスの活用を促す中核的機能と位置付けられ、組織的に知識仲介を行う機関(以下、知識仲介機関 knowledge brokering organisation)<sup>5</sup>であるWhat Works Centre(WWC)の組織や活動内容をみていく。

#### (1) What Works Centreとは

英国のWhat Works Centreは、公共政策に携わる人々、特に地方自治体職員や対人援助職など現場の実践者の意思決定において、何が有効か(what works)に関する最良のエビデンスを提供することを目的とする機関であり(内山ら 2019)、研究知見、特にランダム化比較試験やシステマティック・レビュー等の手法で創出された頑健なエビデンスを業務の中心に位置付けている点が特徴となっている。英国中央政府が推進しているWhat Worksイニシアティブ(主に有効性に関するエビデンスの活用により、公共政策のパフォーマンス向上を目指す取組み)ではWWCが中核的存在と位置付けられ、政府支出のうち総額2,000億ポンドを超える規模の幅広い社会政策にWWCが関連しているとされる。公共政策に頑健性の高いエビデンスを活用するため行政自身が組織を新設する取組みは米国が源流とされるが、英国でのWWCの設置は、より直接的には1999年に英国で設立された医療・ヘルスケア分野のThe National Institute for Health and Care Excellence(NICE)のモデルを社会政策の各分野へ横断的に展開するとの構想による(Cabinet Office 2013)。

WWCは英国内閣府がWhat Worksネットワークの一員として認定する形式を取っており、2022年3月末時点でWhat WorksネットワークにWWCのフルメンバーとして認定されている機関は9団体で、その他にアフィリエイトが3団体、アソシエイトが1団体となっている。各WWCの成立背景はそれぞれ異なるが、英国内閣府では認定基準として6つの要素を提示している(Cabinet Office 2018)。

- ・独立性：利用者に偏りなく助言を提供する
- ・方法論の厳密性：エビデンスの質を確保する
- ・実用性：特定の政策分野でエビデンスの活用や創出に指導的役割を果たす
- ・アクセシビリティ：利用者にとって分かりやすく活用しやすいよう情報提供する
- ・力量形成：実際にエビデンスが活用されるように働きかける
- ・透明性：情報提供する内容の方法論や限界

を分かりやすく示す

WWCのような知識仲介機関は、エビデンスを業務の中心に据えていること、多種多様な経歴を有するスタッフ構成や知識仲介のツールによりネットワーク・活動の幅が広いこと、行政から独立しつつも距離が近いことが特徴である(MacKillop and Downe 2022)。WWCも大学や職能団体、地方自治体等と協働で頑健なエビデンスに関する研究プロジェクトや人材育成プログラム等を実施すると同時に、中央政府との人的・財政的繋がりが強く、省庁の幹部級職員との会合も定期的に設けられるなど信頼関係や影響力を有している。学術研究機関は行政から資金的に独立性を保ちつつ主体的に学術研究の知見を公表しているし、シンクタンクは政策立案者と近い関係性を構築することで政策のアイデアやニーズに応じた知識を提供することに重点が置かれている。これに対してWWCをはじめとした知識仲介機関では、両者の特色を兼ね備えられるようバランスを取りながら運営されており、研究の信頼性や正当性を確保しつつ、行政にとって有用となるようタイムリーかつ理解しやすい形で知識の提供を行うことに努めている(MacKillop and Downe 2022)。

## (2) 基本的機能

WWCの基本的機能は、社会政策として行われる各種取組みの有効性に関するエビデンスの「創出」、利用しやすい形態への「翻訳」、政策・実践への「適用」とされる(Cabinet Office 2018)。このうちエビデンスの創出に関して、WWCでは二次研究としてエビデンスの統合を行うことが重要な機能とされているが、それだけに留まらず、当該政策分野におけるエビデンス基盤(evidence base)において政策や実践の意思決定に不足している知見を特定し、そのギャップを埋めるために、一次研究として形成的評価や総括的評価の推進を担っているWWCもある。また、政策や実践の意思決定に関与しエビデンスを実際に活用する立場にある人々の参画を促して、実務に浸透させようとする試みもみられ(EPPI Centre 2018)、これはエビデンスの創出と適用を同時に図る取組みであると言える。

エビデンスの翻訳の観点では、様々なレベルで公共政策の意思決定にエビデンスが活用されることを目指した取組みが行われている。特に対人サービスに関しては、主に地方自治体の政策担当者やサービス提供部門(公務員だけでなく委託先となる民間団体も含む)の管理職といったメゾレベルにあたる人々が、中央政府(マクロレベル)の政策と現場の個別対応(ミクロレベル)の狭間で、双方の要請に応える効果的・効率的な意思決定ができるよう、WWCでは理解しやすく整理された形式での情報提供方法を開発してきた。この背景として、対人サービスの実務家には省察的实践が求められるが、実態として意思決定は限定合理性の中で行われるため、その実行支援として知識仲介の必要性が広く認識されるようになったことが挙げられる(Boaz et al. 2019)。WWCによる翻訳の代表的な手法としては、システマティック・レビュー等の二次研究によって創出したエビデンスを、エビデンスの頑健性や有効性といった軸で数値化したり、関連する研究を手掛ける英国内の研究機関を地図で示したり、実装方法に関する研究やガイダンスを併記する等、意思決定支援や実装支援を目的としてエビデンスを比較可能な形式にし、ウェブサイト上で可視化したツールキット(Toolkit)の方式が複数のWWCで採用されている。

エビデンスの適用では、WWCが地方自治体の政策担当者やメゾレベル～ミクロレベルの対人援助職へのEBPプログラムの普及や実装の役割を担っている。普及活動や実装支援の内容は多岐にわたっており、広報やワークショップ等にとどまらず、公務員を対象とした研修コースの設置、対人援助職による常設パネルの運営、ステークホルダーが参画する共同研究の実施、あるいはより踏み込んでエビデンスに基づく政策や実践に関する勧告(ガイダンス)を発出するWWCもある。多くのWWCは設立当初から普及活動に注力してきたものの、エビデンスによる政策や実践の意思決定の変化がなかなか進展しなかったとの反省から、近年はより適用に焦点化し、実務家のエビデンスの需要を高めるアプローチを採用するようになってきたとされる

(EPPI Centre 2018)。

### (3) 児童福祉に関連するWhat Works Centre

英国内閣府が運営するWhat Worksネットワークには13団体が加盟しているが、このうち下記の6団体は子ども・若者に関連する政策分野をカバーしており、前半3団体がフルメンバー、後半3団体がアフィリエイトメンバーである。各団体がフォーカスを置く政策分野は近接ないし部分重複しているが、ミッションや当該分野へのアプローチ方法はそれぞれ独自性があり、設立背景や立場も異なる。同じターゲット層へアプローチする場合は協調的に活動しており、共同プロジェクトの実施例も多い<sup>6</sup>。

- ・ Early Intervention Foundation (EIF)
- ・ What Works for Children's Social Care (WWCSC)
- ・ Education Endowment Foundation (EEF)
- ・ Youth Endowment Fund (YEF)
- ・ Youth Futures Foundation
- ・ Centre for Transforming Access and Student Outcomes in Higher Education

このうちWhat Worksネットワークのフルメンバーであり、児童福祉分野と特に関連の深いEIFとWWCSCについて概観する(表1)。

EIFは子ども・若者に関連した早期支援・早期介入プログラムを扱っているWWCで、教育省の補助金を原資に非営利団体として創設された。EIFでは、100件を超える早期介入のEBPプログラムを通覧できるEIFガイドブック(EIF Guidebook)をツールキット方式で作成しており、児童虐待の未然防止、メンタルケア、子ども・若者の非行抑止といった多様なテーマのEBPプログラムを掲載している。子どもやその家庭への早期支援・早期介入は、身体・認知・社会・情動性・行動といった幅広いアウトカム領域に影響があると見込まれ、EIFでは子ども・若者の長期的予後の改善をミッションとして掲げている。

また、エビデンスの確からしさに関して団体としての定義や基準を定めたEIFエビデンス・スタンダード(EIF Evidence Standard)は中央政府でも援用されており、例えばChildren's Social Care Innovation Programmeでは教育省がサービス提供

主体に対して研究計画の作成段階で対照群を設定した評価設計とするよう求める際にEIFエビデンス・スタンダードが用いられた。EIFの主なフォーカスは中央政府の早期支援イニシアティブとして交付される資金を執行する地方自治体の政策担当者の意思決定支援や実装支援であり、EIF自体はプログラムの評価や研究者への資金提供は行っていないが、後述するように、YEF等の他団体と連携して形成的評価やエビデンス基盤の強化に関与している。

WWCSCは児童ソーシャルケアにおけるエビデンスの提供を推進する団体で、教育省が2017年に運営団体を公募し、イノベーション推進をミッションとする政府系の外郭団体であるNestaが立上げ期の委託を受けて設置されたWWCである。非営利団体として独立したのは2020年で、その後新型コロナウイルス感染症の拡大が生じたため、アニュアルレポート等は2022年3月時点で公表が延期された状態である。ソーシャルワーカーをはじめとする対人援助職の参画やネットワーク形成を重視し、重要な活動方針の決定では意見聴取や常設パネルへの諮問等を行ってきた(Alliance for Useful Evidence 2020)。エビデンス創出では4つの原則(研究がもたらす影響、政策分野の複雑性、活用時の実用性、対人援助職のエンパワーメント)を考慮することとし、実証研究や評価研究が行われている。翻訳の観点では、児童ソーシャルケアのEBPプログラムのレビュー結果をウェブサイトにツールキット形式で整理・公表しているほか、Children's Social Care Innovation Programmeで開発が進められた44のプログラムに関する個別の評価報告書(計2,000ページ以上)を約100ページのレポートに要約(What Works for Children's Social Care 2021)するなど、対人援助職の意思決定を支援している。

## 4. 対人サービスのEBPプログラム発展の取組み

児童福祉分野の中でも早期支援・早期介入や児童ソーシャルケアといった予防的取組みの必要性が高い分野では、英国中央政府がWWCの設

表1 児童福祉分野のWhat Works Centre

団体名	Early Intervention Foundation (EIF)	What Works for Children's Social Care (WWCSC)
政策分野	子ども・若者への早期支援・早期介入	児童ソーシャルケア
ビジョン/ ミッション	<ul style="list-style-type: none"> <li>・ビジョン：すべての子どもが十分にポテンシャルを発揮できる</li> <li>・ミッション：効果的な早期介入を実施可能とし、活用されることにより、良くない予後のリスクに晒されている子ども・若者の生活改善を保障する</li> </ul>	セクター全体の意思決定者に、その時点で利用可能な最良のエビデンスを提供し、子ども・若者・家族によりよいアウトカムをもたらす
設立	2013年7月、教育省の補助金により非営利団体として設立。直接的な契機は早期介入に関する2011年の政府報告書	2017年10月に教育省がNestaを中心としたコンソーシアムに設立及びインキュベーションを委託し、2020年4月から非営利団体として独立
年間支出額 (千ポンド)	<ul style="list-style-type: none"> <li>・2018年度：1,725千ポンド</li> <li>・2019年度：2,564千ポンド</li> <li>・2020年度：3,035千ポンド</li> </ul>	(アニュアルレポート未公表)
資金拠出元	約90%が中央省庁(教育省、労働年金省、レベリングアップ・住宅・コミュニティ省、内務省)	教育省
主要な アウトカム	子どもの発達に関する以下の事項： <ul style="list-style-type: none"> <li>・身体：肥満、心身の健康</li> <li>・認知：学業達成、労働参加</li> <li>・社会性・情動：精神的健康、ウェルビーイング</li> <li>・行動：犯罪・暴力、反社会的行動</li> </ul>	<ul style="list-style-type: none"> <li>・子ども・親・養育者・家族の権利</li> <li>・子ども・若者のアウトカム(客観及び主観)</li> <li>・親・養育者・家族のアウトカム</li> <li>・中間アウトカム：組織的要因(費用対効果、労働力アウトカム、対人援助職のスキル・知識・経験)</li> </ul>
エビデンスの 基準	EIFエビデンス・スタンダード <sup>7</sup> と呼ばれる、プログラムの有効性に関する実証状況を5段階に分類する枠組みを開発・使用	エビデンスの厳密性に関してはEIFの基準を準用しつつ、有効性だけでなく実装方法や文脈的要因にも着眼するため、EMMIEフレームワーク <sup>8</sup> を採用
「創出」の 主な取組み	早期介入に関する既存の評価研究等を収集・整理しレポートを公表するほか、評価方法のガイドラインも作成。なお、EIF自身がプログラム評価や一次研究への資金提供を直接手掛けることはない	子ども・養育者・対人援助職等のステークホルダーの意見を踏まえ、研究の優先順位を決定。実証研究や評価研究を実施したり、開発段階のプログラムに研究助成を行ったりしている
「翻訳」の 主な取組み	国内外で開発された100以上の早期介入プログラムの評価レポートをレビューし、スタンダードに基づきエビデンスのレベルやコストを点数化した上でEIFガイドブックとしてウェブサイト上で公表。子どものアウトカムや年齢層で検索・並べ替えが可能	国内外の児童ソーシャルケアに関するプログラムのレビュー結果をエビデンス・ストアとして集約して公表。アウトカム別に有効性やエビデンスの強さをレーティング。実装に影響する要因や英国国内での実施状況も紹介
「適用」の 主な取組み	地方自治体やサービス提供主体と連携し、早期支援プログラムの実装が地域のニーズや資源に適した形となるよう、情報・助言の提供や知見交流を促進。また、エビデンス基盤が充実しているテーマでは、推奨事項をガイダンスとして発行	スクールソーシャルワーク、対人援助職のウェルビーイングといった個別テーマで、自治体と共同研究を実施。また、ソーシャルワークの実践に研究を取り入れる実践的方法や、研究の批判的吟味及び適用方法を学ぶ1-2日間の研修コースを開催

(出所) Early Intervention Foundationウェブサイト及びWhat Works for Children's Social Careウェブサイトより筆者作成

置を主導してきた。WWCはエビデンスの基準に従って既存の研究をレビューし、ツールキット等の分かりやすい形に翻訳するほか、実装段階での情報提供や助言等を通じて地方自治体の政策担当者や現場の対人援助職によるEBPプログラムの適用を促してきた。この過程で、WWCは当該分野におけるエビデンス基盤と政策や実践とのギャップも特定でき、その洞察はエビデンスの創出にも有用である。ここでは、WWCが形成的評価に関わることで、イノベーションをEBPプログラムへと段階的に発展させる取組みを紹介

する。

### (1) EIFとYouth Endowment Foundationの協働によるEBPプログラム開発

早期介入をテーマとするEIFは、自ら評価研究によるエビデンスの創出を行う機能は有していないものの、広く用いられているEIFエビデンス・スタンダードと連動した10段階の評価プロセス(10ステップ)を提示しており、形成的評価によるEBPプログラムの発展を促進している(表2)。これを全面的に取り入れたのが、WWC

のアフィリエイトとして2018年10月に新設されたYouth Endowment Foundation (YEF) である。YEFは内務省による2億ポンドの助成金と民間助成財団の研究助成により10年間を目途にエビデンス創出及びその普及活動を行っており、EIFはその運営にも参画している。

YEFが資金助成プログラムの評価で用いる評価者向けガイドも10ステップに沿って作成しており、実現可能性を調査するプロジェクトであれば4段階目、パイロット検証であれば5段階目までが完了するような評価設計が求められる。YEFの資金助成プログラムの一例として、A supportive homeと称する資金助成ラウンドでは5～10程度のプログラムに総額600～1,000万ポンドを実証及び評価に提供する予定であるが、10ステップでの6段階以上、または2年以内に5段階目まで進めるプログラムのみが資金助成を受けられる対象とされている (Youth Endowment Fund 2021a)。またYEFは、EIFへエビデンスレビューを委託することで、ペアレンティングプログラムや家庭内暴力・虐待への介入といった、有効性が有望視されているものの英国国内での実証結果が乏しい家庭支援プログラムのエビデンス基盤を総括し、資金助成を検討する主体に情報提供も行っている (Youth Endowment Fund 2021b)。EIFの知見を基礎として、YEFでは他にも助成申請希望者へ10ステップに沿った形成的評価の実施方法を解説するワークショップを開催したり、

表2 EIFの評価の10ステップ

段階	位置付け	EIF エビデンス・スタンダードにおけるレベル
1	セオリー・オブ・チェンジの確認	NL2
2	ロジックモデルの構築	
3	実行計画の作成	
4	実現可能性の調査	
5	アウトカムのパイロット検証	レベル 2
6	効用 (efficacy) の実証	レベル 3
7	有効性 (effectiveness) の実証	レベル 4
8	精緻化・モニタリング	
9	適応・移転	
10	規模拡大	

(出所) Early Intervention Foundation (2019) p.6より。邦訳は筆者による。

評価設計の段階で外部評価者の伴走支援を提供したりすることで、EBPプログラムの開発に寄与している。

## (2) WWCSによる評価伴走プログラム：Practice In Need of Evidence

WWCSはイノベーションの推進においてステークホルダーの参画を重視しており、地方自治体やサービス提供主体等の実務家から将来的な拡大が期待できる対人サービスの実践方法の提案を募集し、評価を通じて伴走支援するプログラムを実施している (What Works for Children's Social Care Practice In Need Evidence プログラムウェブサイト)。このプログラムはPractice In Need of Evidence (PINE) と呼ばれ、2022年2月から3ラウンド目の申請を募集中である。現在のラウンドは5つの優先テーマ (児童ソーシャルケアの担い手、子ども・若者のメンタルヘルス、施設養育、家庭内暴力、リスクと意思決定) が設定されており、採択された場合、申請者は最長18か月間にわたり最大で総額35万ポンドのプロジェクト実施資金を得ることができる。

PINEは申請プロセス自体が評価計画を設計する内容になっており、申請時の4つの段階 (Describe: ロジックモデル等の作成、Plan: データ収集や倫理的側面の検討、Do: 調査・分析、Review: 結果報告とデータの可視化) ごとにWWCSのPINE運営チーム (ソーシャルワーカーや研究者) とのディスカッション機会も設けられている。さらに、申請者は申請書提出後、Readiness for Trial Trainingと呼ばれる実証研究の研修受講が義務付けられており、ランダム化比較試験等の基礎的知識を学ぶほか、採択プロジェクトには独立した評価者としてWWCSの指定研究機関が外部評価を行うことになる。このようにPINEには、形成的評価と評価キャパシティ形成の要素が盛り込まれている。

PINEに採択されたプロジェクトの例として、ロンドン中心部のサザークで実施している父親を対象としたペアレンティングプログラム、Fathers Parenting Groupがある。このプログラムは、既に早期支援の一環として2017年から取組まれており、計10週のグループワークを行う内容で

ある。WWCSCとしては、父親の子育て参画を進めるためのエビデンス基盤の充実を期待しており、PINEにより採択されて2021年1月からWWCSCの伴走支援を受けて評価を実施している。申請書段階では先行研究に言及するのみであった状態から、ロジックモデル作成へと繋がったほか、介入の有効性を評価するための問い（対照群と比較する際の視点）を投げ掛けたり、エビデンスに関する研修コースを提案したりしている。

## 5. おわりに：対人援助においてエビデンスに基づく政策や実践を推進する上での論点

1990年代後半以降の英国では人的資本への投資、NPM、緊縮財政といった政府方針と並行して、中央政府が主導するWhat Worksイニシアティブとしてエビデンスに基づく政策・実践が推進されてきた。児童福祉やその近接分野では早期支援・早期介入といった予防的政策への期待が一貫して高く、対人サービスでのイノベーションを推進する取組みがなされてきたが、ここではクライアントのアウトカムを向上させるEBPプログラムの発展が重視され、様々な政府イニシアティブを通じたEBPプログラムの形成的評価が積極的に実施されてきた。これにより英国の児童福祉分野におけるエビデンス基盤は増強されたが、資金不足や管理主義の影響もあって自治体職員や対人援助職のエビデンスの活用状況は依然、大きな改善の余地を残したままである。こうした環境下で進められてきた児童福祉改革では、特に近年はエビデンスを政策や実践へ活用することへの傾きが強まっており、研究を実務家へ橋渡しする役割を担う知識仲介機関として、WWCが担う役割も拡大している。

英国のこうした実情を踏まえ、日本国内でエビデンスに基づく実践を一層推進する際の今後の論点として、第一に、中央政府だけで実務家にエビデンスの活用を促すことには限界があり、それを乗り越えるための方策が求められる。対人サービスが公共政策として広範囲に提供され

るとしても、実務家はクライアントの個別性に応じつつ、各々の現場における資源・時間・情報等の制約の中で意思決定を行うため、中央政府のトップダウンによる画一的なアプローチは基本的に馴染みづらい。そこで、行政から独立した立場で頑健なエビデンスを蓄積し社会的信用を得てきた学術研究機関と、行政のニーズを理解し知識を実用的な形態で提供することに長けたシンクタンクの、双方の特性を有しながらも、さらにネットワーク形成に優れたWWCのような知識仲介機関のみが果たしうる役割もあるのではないかと。諸外国で取組まれてきた知識仲介機関の実例や研究に学びつつ、日本の政策や実践の文脈に沿った検討が望まれる。

第二に、質の高いエビデンスが政策や実践へのエビデンスの活用につながるわけではなく、エビデンスを活用する実務家も考慮した取組みを展開する必要がある（小林 2020）。英国の児童福祉分野の予防的政策でも、中央政府における検討過程ではエビデンスの存在が根拠とされていたものの、実装段階では政策や実践とエビデンスが結びつかなくなる例がみられる。こうした研究・エビデンスと実際の政策・実践との間にあるギャップを解消する方策として、英国ではWWCのエビデンスの創出・翻訳・適用といった本来機能の発揮に加えて、組織運営や意思決定に実務家の参画を求めたり、既存のエビデンス基盤を俯瞰的にみた上で研究資金を実務に接続しやすい領域へ焦点化したりする工夫を講じている。対人サービスの実態に即しつつエビデンスに基づく政策・実践を推進するには、現場で重視されている臨床的知識・経験と研究によるエビデンスを融合させるという難問と向き合う必要があるが、知識仲介機関ではスタッフ構成に実務経験者も含めて多様性を確保していることも取組みを進める上での参考になるだろう。

第三に、対人サービスにおけるイノベーションをEBPプログラムへと発展させる手法としての形成的評価を、エビデンスに基づく政策・実践を推進する一連の取組みといかに関連付けられるかが重要だと考えられる。英国の児童福祉分野の事例では、WWCが形成的評価と総括的評価の両方を包含するガイドラインを作成し、中央

政府が研究資金提供の一環としてこのガイドラインを採用することで形成的評価の実施を要件化し、その実施に際して伴走支援等の評価キャパシティ形成が提供されていた。対人サービスのようなエッセンシャル・ワークにおける評価では、効率・節約を規準とした業績評価の偏重がみられるが、業務改善やサービス品質向上につながるよう、安心・安全を規準としたプログラム評価の実施が望まれている（藤井 2022）。日本国内でも、対人援助職による革新的な試みが形成的評価を通じてEBPプログラムとして結実しやすくなるよう、研究・政策・実践に携わるステークホルダー間での本格的な議論がまたれる。

## 注記

- 1 対人サービスは、政策と実践の局面を担う主体が行政担当者から対人援助職に移行する点に特徴がある。実態としては、円環状の政策サイクル論のようなシンプルで一貫した連続性は確保されておらず、政策と実践の間で複雑に相互作用が生じるが、本稿では便宜上、政策と実践に局面を分けて整理している。
  - 2 用語の明確な定義はなされていないが、エビデンス情報に照らした実践（evidence-informed practice）という表現によって対人援助職の抵抗感は薄まりつつあるとされる。
  - 3 シュア・スタートのほか、8～13歳の脆弱な立場にある子どもに地域の民間団体から支援プログラムを提供するChildren's Fund、13～19歳を対象とした普遍的サービスとして導入されたThe Connexions等のイニシアティブも同時並行で開始された。これらの予防的なプログラムには、2002年度だけで合計8億100万ポンドもの資金が投じられている。
  - 4 Early Intervention Foundation（2020b）によれば、主な対象別に以下の4種類の異なるアプローチへとプログラムが分岐しているとされる。(a) シュア・スタートアプローチ：2013年の法定ガイダンスに沿ったすべての家庭への総合的な幼児サービス (b) Early Helpアプローチ：ハイニーズ家庭に焦点化した全年齢層の脆弱な子ども家庭への早期支援モデル (c) 公衆衛生的アプローチ：乳幼児を主な対象とする保健分野との連携を強化 (d) 就学準備アプロ
- チ：幼児教育や小学校とより密接に連携しており、運営者も学校法人であることが多い。
- 5 知識仲介者（knowledge broker）が個人または機関を指し、取り扱う情報や財源等により活動形態が多様だとされる（ナトリール 2015）のに対し、WWC等の知識仲介機関は独立した組織として公共政策のステークホルダーに研究の知見を提供する（MacKillop and Downe 2022）。また、教育分野に関する先行研究ではエビデンスの集約や伝達を専門的に担う機関をエビデンス仲介機関としている（森・岡崎 2020）が、児童福祉分野では臨床的知識・経験も重要と位置付けられるほか、WWCによる評価キャパシティの形成等も重要な機能となるため、本稿ではエビデンスを業務の中心としつつも幅広い研究知見を実務家へと橋渡しするという趣旨から、WWCを知識仲介機関と呼称している。
  - 6 例えば、EIFとEEFはエビデンス基盤と実践のギャップ解消に向けてWWCが今後どのようにアプローチすべきかを検討するため、各々が取組んできた社会情動的学習を題材に行動科学の知見から知識移転を図る方法を考察し、共同でレポートをまとめている（Early Intervention Foundation 2020b）。また、EIF、WWCSC、Centre for Transforming Access and Student Outcomes in Higher Educationの3団体は、The Evidence Quarterと呼ばれるオフィスに入居しており、物理的にも経営資源やアイデアを共有して社会的影響を高めることが企図されている。
  - 7 EIFエビデンス・スタンダードでは、最上位にあたるレベル4は複数の厳密な評価を通じて長期的な有効性が実証されているプログラム、レベル3は1つ以上の厳密な短期的評価でポジティブな結果があるプログラム、レベル2は有効性に関する予備的エビデンスはあるが因果関係は未検証のプログラム、NL2（レベル2に至っていない状態）はレベル2の要件を満たさないプログラム、NE（有効性がみられなかった）はレベル3相当の実証の結果としてプラスの有効性が検出されなかったプログラム、となる。
  - 8 EMMIEフレームワークは、Effect（介入の有効性）・Mechanisms（介入が機能するメカニズム）・Moderators（機能に影響しうる文脈等の要素）・Implementation（実装方法）・Economic impact（経済的影響）の頭文字を取った名称とされているように、有効性と実装方法や文脈的要因を考慮するための枠

組みである。

## 参考文献

- アイリーン・ムンロー (2021) 『子ども虐待 保護から早期支援への転換 児童家庭ソーシャルワーカーの質的向上をめざして』、明石書店
- 家子直幸・小林庸平・松岡夏子・西尾真治 (2016) 「エビデンスで変わる政策形成—イギリスにおける「エビデンスに基づく政策」の動向、ランダム化比較試験による実証、及び日本への示唆—」、『政策研究レポート』、三菱UFJリサーチ&コンサルティング
- 内山融・小林庸平・田口杜輔・小池孝英 (2018) 「英国におけるエビデンスに基づく政策形成と日本への示唆—エビデンスの「需要」と「供給」に着目した分析—」、『RIETI Policy Discussion Paper Series 18-P-018』、独立行政法人経済産業研究所
- 大島巖・源由理子・山野則子・費川信幸・新藤健太・平岡公一 (2019) 『実践家参画型エンパワメント評価の理論と方法—CD-TEP法：協働によるEBP効果モデルの構築』、日本評論社
- 小池拓自 (2020) 「エビデンス仲介機関としての英国WWCLEGの取組—英国における地域経済成長政策とEBPM—」、『国立国会図書館 調査及び立法考査局』、835:1-28
- 小林庸平 (2020) 「日本におけるエビデンスに基づく政策形成 (EBPM) の現状と課題—Evidence-Basedが先行する分野から何を学び何を乗り越える必要があるのか—」、『日本評価研究』、20 (2) :33-48
- サンドラ・M.ナトリー、イザベル・ウォルター、ヒュー・T.O.デイヴィス (2015) 『研究活用の政策学 社会研究とエビデンス』、明石書店
- 田邊泰美 (2019) 『現代イギリスの児童虐待防止とソーシャルワーク 新労働党政権下の子ども社会投資・児童社会サービス改革・虐待死亡事件を検証する』、明石書店
- 藤井誠一郎 (2022) 「地方自治体のエッセンシャル・ワーク評価」、『日本評価研究』、22 (1) : 29-42
- 森俊郎・岡崎善弘 (2020) 「教育分野におけるエビデンス仲介機関の特徴」、『日本評価研究』、20 (2) : 77-88
- Alliance for Useful Evidence (2020). *A Practical Guide for Establishing an Evidence Centre.*
- Boaz, A. Davis, H. Fraser, A. and Nutley, S. (2019). *What Works Now? Evidence-informed policy and Practice.* Policy Press.
- Cabinet Office (2013). *What Works: evidence centres for social policy.*
- Cabinet Office (2018). *What Works Network :Membership Criteria.*
- Cairney, P. and Denny, E. (2020). *Why Isn't Government Policy More Preventive?.* Oxford University Press.
- Department for Education (2011). *The Munro Review of Child Protection: Final Report A child-centred system.*
- Department for Education (2017). *Children's Social Care Innovation Programme: Final evaluation report.*
- Department for Education (2020). *Children's Social Care Innovation Programme: Round 2 Final report.*
- Early Intervention Foundation ウェブサイト  
<https://www.eif.org.uk/> (2022年3月31日閲覧)
- Early Intervention Foundation (2019). *10 steps for evaluation success.*
- Early Intervention Foundation (2020a). *Planning early childhood services in 2020: Learning from practice and research on children's centres and family hubs.*
- Early Intervention Foundation (2020b). *Developing a behavioural approach to knowledge mobilization: Reflections for the What Works Network.*
- Eisenstadt, N. Oppenheim, C. (2019). *Parents, Poverty and the State.* Policy Press
- EPPI-Centre (2018). *UK What Works Centres: Aims, methods and contexts.*
- MacKillop, E. and Downe, J. (2022). Knowledge brokering organisations; a new way of governing evidence. *Evidence & Policy*, Early View (2022年3月31日閲覧)
- What Works for Children's Social Care ウェブサイト  
<https://whatworks-csc.org.uk/> (2022年3月31日閲覧)
- What Works for Children's Social Care Practice In Need Evidence プログラムウェブサイト  
<https://whatworks-csc.org.uk/working-with-practice/practice-in-need-of-evidence/> (2022年3月31日閲覧)
- What Works for Children's Social Care (2021). *Learning from innovation: Summaries of the Children's Social Care Innovation Programme round 2 evaluations.*
- What Works Network (2018). *The What Works Network: Five Years On.*

Youth Endowment Fund (2021a). *A supportive home: helping families to overcome challenges.*

Youth Endowment Fund (2021b). *Supporting families, preventing violence: the evidence base.*

(2022.6.28 受理)

## **UK initiatives to develop innovations in human services into Evidence-Based Programme －Focus on the What Works Centres in the child welfare－**

Naoyuki Ieko

Mitsubishi UFJ Research & Consulting  
ieko@murc.jp

### **Abstract**

In the UK, the central government has promoted a policy of emphasising evidence in social policy since around 2000, and evidence-based practice has been required for human services. Evidence-based initiatives in the field of child welfare have been promoted in conjunction with reforms to improve the professionalism of human service workers, and there are particularly high expectations for preventive policies such as early help and early intervention. As austerity has shifted the policy focus from large-scale trial to the use of evidence. Central government-led What Works Centre has been established, primarily as a knowledge brokering organisation to support the use of evidence on effectiveness in policy and practice decision-making. The function is to generate, translate and apply evidence. In the child welfare sector, the approach is linked to the funding of formative evaluation to develop innovations in human services into Evidence-Based programmes.

### **Keywords**

Human services, evidence-based policy and practice, preventive policy,  
knowledge brokering organisation, formative evaluation



## 【実践・調査報告】

# 現場実践を効果モデルに発展させる評価ガイドの開発と実装 — CD-TEP 法を活用した事例の考察 —

清水 潤子

武蔵野大学<sup>1</sup>

jshimizu@musashino-u.ac.jp

新藤 健太

日本社会事業大学

shindo@jcsu.ac.jp

## 要 約

本稿は、「プログラム理論 (Theory:T)」と「エビデンス (Evidence:E)」と「実践 (Practice:P)」との「円環的対話 (Circular Dialogue: CD)」(CD-TEP法)を活用し作成した障害者芸術文化活動普及支援ガイドの事例について振り返り、CD-TEP法の現場における形成的評価の促進・定着と、実装可能性について分析・考察を行ったものである。その結果、評価ガイドの開発や活用支援により、普及支援事業の支援センターが評価を事業改善に活かす等、形成的評価の現場における普及がうかがえたが、一方で評価ガイドにおける効果モデルやフィデリティ尺度の役割が薄まるという現象が見られた。また実装科学のフレームワーク(CFIR)を部分的に用いてCD-TEP法の実装性について考察した結果、評価アプローチとしての理念を実践に浸透させるためには、様々な実装上の課題を解決するための体制や仕組みが必要であり、実装を支える技術的支援の重要性が明らかとなった。

## キーワード

形成的評価、CD-TEP、評価ガイド、中間支援、実装科学

### 1. はじめに

#### (1) 本稿の目的

本稿は厚生労働省の「障害者芸術文化活動普及支援事業（以下、普及支援事業）」において、「プログラム理論 (Theory:T)」と「エビデンス (Evidence:E)」と「実践 (Practice:P)」との「円環的対話 (Circular Dialogue: CD)」(以下、CD-TEP法<sup>2</sup>) (大島 2011)を活用して作成した「障害者芸術文化活動普及支援ガイド<sup>3</sup> (以下、評価ガイド)」の開発・実装事例を振り返る。そして、①CD-TEP法は現場における形成的評価の促進・定着に寄与するかという観点と、②CD-TEP法の

実装可能性という2つの点について分析・考察し、実践において効果的であった点や課題について提示するものである。

#### (2) 障害者芸術文化活動普及支援事業評価ガイドライン作成事業

##### ①障害者芸術文化活動普及支援事業

厚生労働省では「障害者の芸術活動普及支援モデル事業（以下、モデル事業）」(平成26年度～平成28年度)および、「普及支援事業」(平成29年度～)により、障害のある人たちの芸術文化活動の振興を図るとともに、自立と社会参加を促進するための地域における支援体制を全国に展

開してきた（令和2年度障害者芸術文化活動普及支援事業連携事務局 2021）。

普及支援事業の事業展開イメージは図1の通りであり、都道府県「障害者芸術文化活動支援センター（以下、支援センター）」、ブロック「障害者芸術文化活動広域支援センター（以下、広域センター）」、全国「連携事務局」といった支援拠点が設置されている。

支援センターは都道府県の支援拠点として、人材育成や相談支援、関係者のネットワークづくり、発表機会の確保、情報収集等の多様な支援事業を展開している。令和2年度は35都府県、37の支援センターが設置された。

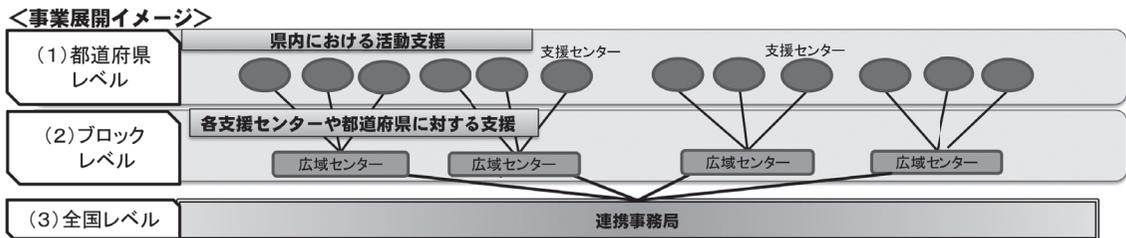
広域センターは、全国を7つの地域別ブロックに分け、ブロック内の支援センターのサポート

や、支援センターが未設置地域の事業所等に支援事業を展開している。

支援センター、広域センターを横断的にサポートする事務局として、「連携事務局」が設置されている。美術分野と舞台芸術分野において、それぞれ1団体ずつ活動している。

このように、都道府県レベルで現場団体（例：障害者施設、教育施設、芸術文化施設等）や当事者をサポートする支援センターと、地域ブロックレベルで支援センターに対して支援を行う広域センターと、専門分野ごとに主に全国の広域センターに対して支援を行う連携事務局が存在し、中間支援組織<sup>4</sup>が三層化された事業運営のシステムとなっている。

図1 障害者芸術文化活動普及支援事業の事業展開イメージ



(出所)厚生労働省「障害者芸術文化活動普及支援事業」(2021)より抜粋

## ②評価ガイドライン作成事業

障害者芸術文化活動の普及支援に関する評価は、モデル事業や平成29年度の普及支援事業実施時までは、評価委員会を設け、事業の成果や課題について評価が行われてきた。しかし、平成30年度の普及支援事業からは、都道府県レベルにおける支援センターの事業に関しては、各都道府県が評価の実施主体となり、評価委員会での評価が行われなくなった。そのため、各自治体が自らの評価等に活用する新たな指標や評価の枠組みが必要となった（日本ファンディング協会 2019b）。このような背景に基づき、「障害者芸術文化活動普及支援事業評価ガイドライン作成事業（採択名）（以下、評価ガイド作成事業）」が実施されることとなり、全国の支援センターや自治体等が活用可能な事業評価ガイドの作成に至った。

## 2. 実践報告の枠組み・方法

本稿は筆者が評価ガイドの開発・活用・改訂に関わった平成30年4月から令和3年3月までの期間を対象とする。各フェーズにおける筆者の役割を表1に示す。

本稿では、まず各フェーズで行われた実践とその結果について明示する。そして、そこから本稿の目的に沿って事例の考察を行う。CD-TEP法の実装可能性についての考察は、実装研究のための総合フレームワーク（Consolidated Framework for Implementation Research：CFIR<sup>5</sup>）の一部を枠組みとして使用する。

評価ガイドの開発・活用・改訂における筆者の位置づけについては、開発期は「平成30年度厚生労働省障害者総合福祉推進事業」に位置づけられた「評価ガイド作成事業」を特定非営利

表1 評価ガイド開発・実装のプロセスと筆者の役割

フェーズ	期間	依頼者	筆者の役割
開発期	平成30年3月～ 平成31年3月	厚生労働省	【評価ガイドの開発】 ・ロジックモデルづくりのワークショップ（ファシリテーション） ・ロジックモデルの整理・構造化 ・支援センター等へのヒアリング ・アウトカム指標とフィデリティ尺度の作成 ・効果的援助要素とアウトカムの関係の検証
活用期	平成31年4月～ 令和3年3月	連携事務局（美術分野）	・フィデリティ尺度の調査・分析 ・支援センターを対象とした評価ガイド活用の研修 ・広域センターを対象とした評価ガイド活用の研修 ・支援センター、広域センター間の学びや気づきの共有に向けたファシリテーション
	令和2年1月～ 令和3年3月	広域センター	・支援センターを対象とした評価ガイド活用の研修 ・支援センター間の学びや気づきの共有に向けたファシリテーション
改訂期	令和2年12月～ 令和3年3月	連携事務局（美術分野）	・支援センターや広域センターのフィードバックをもとにした評価ガイド改訂版の作成

(出所)筆者作成

活動法人日本ファンドレイジング協会が受託し、筆者を中心に作成を行った。

活用期については、普及支援事業の連携事務局（美術分野）からの依頼に基づき、全国の支援センターや広域センターが評価ガイドを活用し、事業改善に役立てられるようになることを目的とした研修を実施した。研修を経て、評価ガイドの意義を感じた一部の広域センターから、同センターが管轄する支援センターに対して、より詳細な研修を実施してほしいという依頼があった。そのため、活用期の半ば以降は広域センター独自の研修依頼にも応じた。

改訂期においては、活用期に受けた評価ガイドに対するフィードバックをもとに、評価ガイドの改訂版を連携事務局の依頼に基づき作成した。

普及支援事業の評価ガイドは、目的に応じて普及支援事業の支援センター、広域センター、連携事務局、実施・未実施自治体の行政が活用できるものとして作成されたが、本稿においては、主に支援センターにおける実践を振り返る。

### 3. 評価ガイドの開発

#### (1) 評価ガイド作成の方向性

具体的な評価ガイドの開発に先んじて行った、モデル事業・普及支援事業で評価委員を担当した有識者及び障害者芸術文化活動に係る専門家（評価ガイド作成にあたり、検討委員として関与）

との意見交換の中で、以下の4点が評価ガイドの開発の方針を特徴づけるポイントとして明示された。

#### ①結果の評価から成果の評価へ

これまでは「何を何回実施した」「何に何人参加した」というような結果の評価が中心であった。しかし、それだけでは活動や事業によって制度・施策の目標の達成に近づき、成果を生み出すことが出来たかを評価することが困難なため、成果評価の実施が評価のスコープに含まれた。

#### ②事業改善に役立つ評価

評価の目的は多様である一方で、一般的には「第三者」による「事後評価」や「説明責任」が評価のイメージと併存していることが多い。そのため、このような評価に対する誤解や現場の抵抗感をぬぐいながら、事業を通じて成果を高めていく必要がある。そして、自治体や支援センターが評価を活用し、活動からの学びを事業改善に活かし、普及支援事業全体によりよい効果をもたらすことが重視された。

#### ③現場や関係者とともに実施する協働型評価の導入

普及支援事業の実施主体は平成30年度より自治体となり、民間団体への委託や補助という形式をとる中で、自治体と、支援センターという特徴の異なる2つの組織が主な事業評価の実施主

体となった。特に中間支援組織である支援センターは、自治体に限らず、当事者や施設、教育機関や地域住民といった多様なステークホルダーを抱えており、多様な主体の価値観や視点を反映する際に、共通言語となるような評価ガイドが求められた。

#### ④全国展開に向けた効果モデルづくり

評価手法の未整備と相まって、普及支援事業が制度・施策の趣旨から全国で展開される際に、どのような中間支援組織の活動が効果を生んでいるのかという情報の整理も必要であった。元来、支援センターの多くは、障害者を支援する施設において芸術文化活動支援を実施してきた「プレイヤー」であり、先進性や独自性が際立つ一方で、ネットワークの紡ぎ方等が属人的であるという課題もあった。

また、どのような中間支援体制が整うことで、より障害当事者の芸術文化活動の参加が促進されるのかといった情報は、特に未実施自治体での制度化や予算化において期待される部分があった。そして、自治体が事業を開始したところで、他の自治体では先進事例がある一方、情報の未整備によって「何から手を付けて良いのかわからない」という状況があることも散見された。そのため、地域に成果を生み出している中間支援現場の工夫やイノベーションからの学びを効果モデルとして構築し、全国で活用できるものが期待された。

#### (2) 評価ガイドのデザイン

前項の①と②で提示したポイントを反映した評価ガイドを開発するために、まずわれわれはプログラム評価の枠組みを採用することとした。大島ほか（2019:3）はプログラム評価の導入の対象を「社会問題を緩和し社会状況を改善するという明確なゴール・目標をもち社会的に導入された社会プログラム」としており、普及支援事業は前述の通り厚生労働省が主管する事業であり、「芸術文化活動を通じた障害者の自立と社会参加」という制度の目的や理念も明らかであった。プログラム評価の目的は、形成的評価と総括的評価の2つに大別されるが、特に②の事業改

善に役立つ評価の観点は、「プログラムの形成・改善・改良を導くために、評価情報を得て活用することを意図した評価目的（大島ほか 2019、p.7）」をもつ形成的評価と整合した。

また③や④のポイントについては、多様なステークホルダーとともに、成果を生み出す効果的プログラムモデルを開発・発展させ、実施や普及、改善を多様なステークホルダーとともに行うCD-TEP法に基づいて作成した。CD-TEP法は大島（2011）によって開発され、「プログラム理論（Theory:T）」と「エビデンス（Evidence:E）」と「実践（Practice:P）」との「円環的対話（Circular Dialogue:CD）」に基づいて、社会プログラムに関わる関係者が協働で、「EBP効果モデル」構築を目指す評価アプローチ法である。

#### (3) 開発プロセス

CD-TEP法は大きく3つのステージにより構成されており、それぞれ「Ⅰ. プログラム開発評価・評価基盤形成ステージ」、「Ⅱ. 効果的プログラムモデルへの発展評価ステージ」、「Ⅲ. 効果的モデルの実施・普及・更新評価ステージ」とされる（大島 2011）。そして、第Ⅰステージには3つ、第Ⅱステージには11、第Ⅲステージには4つ、計18のプログラム評価が直面する課題が「課題プロセス」としてまとめられている。CD-TEP法の課題プロセスと評価ガイド作成事業が実施した内容の対比は、表2に記した通りである。「Ⅰ. プログラム開発評価・評価基盤形成ステージ」は、それまでのモデル事業や普及支援事業の成果報告書の調査分析からの情報収集や、厚生労働省との意見交換によって進められた。

「Ⅱ. 効果的プログラムモデルへの発展評価ステージ」から「Ⅲ. 効果的モデルの実施・普及・更新評価ステージ」の「Ⅲ1-2. 効果的プログラムモデルの実施・普及モデル、実施ツールキットの作成」においては、ロジックモデル（表3）と効果的援助要素<sup>6</sup>について、どのように活動の改善と効果モデルの普及に役立つか、という観点を重視しながら評価ガイドの作成を行った。

特徴的なこととして、特に現場には専門用語がなじまないということもあり、CD-TEP法における「効果的援助要素」は「活動のコツ」とい

表2 CD-TEPの課題プロセスと本事業の対比<sup>7</sup>

	プロセス課題	本事業で実施した範囲
Ⅰ. プログラム開発評価・評価基盤形成ステージ		
I 1-1.	ニーズ把握とプログラムゴール・標的集団の設定	○
I 2-1.	既存・試行プログラムの現状把握	○
I 2-2.	プログラム評価可能性・再編可能性アセスメントの実施	○
Ⅱ. 効果的プログラムモデルへの発展評価ステージ		
II 1-1.	プログラム理論の構築・再構築：インバウト理論	○
II 1-2.	プログラム理論の構築・再構築：プロセス理論（サービス利用計画）	—
II 1-3.	プログラム理論の構築・再構築：プロセス理論（組織計画）	—
II 2-1.	効果的援助要素リストの作成	○
II 2-2.	効果的プログラムモデルの実施マニュアルの作成	△
II 3-1.	アウトカム評価尺度・指標の設定と活用計画	○
II 3-2.	アウトカム評価調査の実施とその評価結果の活用	○
II 3-3.	効果的プログラムモデルのフィデリティ尺度作成と活用計画	○
II 3-4.	フィデリティ評価調査の実施と評価結果の活用	○
II 3-5.	アウトカム評価とフィデリティ評価・効果的援助要素の関連性の検証、評価結果の活用	○
II 4-1.	効果的プログラムモデルの構築	○
Ⅲ. 効果的モデルの実施・普及・更新評価ステージ		
III 1-1.	効果的プログラムモデルの実現場への移転、実践的適合モデルの作成	○
III 1-2.	効果的プログラムモデルの実施・普及モデル、実施ツールキットの作成	○
III 1-3.	効果的プログラムモデルの制度化、制度モデルの作成	—
III 2-1.	実践現場の創意・工夫と基盤にした効果的プログラムモデルの改訂・更新	—

(出所)大島(2011)をもとに筆者作成

表3 評価ガイドのロジックモデル

事業の作戦体系図（ロジックモデル）	
最終アウトカム	芸術文化活動を通して、障害者が社会参加できるようになる
中間アウトカム	芸術文化活動を通して、障害者が個性と能力を発揮できる
01	より多くの障害者が芸術文化活動に取り組むことができる
0101	地域の障害者に関する芸術文化活動の現状が正しく理解される 010101 地域の障害者の芸術文化活動について網羅的なアンケート調査を実施する 010102 地域の障害者の芸術文化活動について訪問による詳細な調査を実施する 010103 障害者の芸術文化活動に関する既存の調査結果を参照する
0102	障害者の芸術文化活動に関する相談事が解決される 010201 支援センターに障害者が相談しやすい環境を整備する 010202 多様な通信手段で相談を受け付ける 010203 相談の内容に応じて、関係機関の紹介を行う
0103	障害に配慮した芸術文化活動の情報が収集・発信され、それが届く 010301 様々な機関・団体から芸術文化活動に関する情報を得る 010302 障害者にも分かりやすい伝え方で、芸術文化活動についての情報提供を行う
02	障害者が芸術文化活動に主体的に参加できる機会が増える
0201	障害者が作品を創造し、発表する機会が増える 020101 障害者の作品や障害のある芸術家を調査・発掘する 020102 地域の障害者による作品や芸術文化活動のアーカイブ（記録及び保存）のための活動を行う 020103 障害種別を問わず、参加可能な芸術文化活動の環境整備を支援する 020104 展覧会や舞台公演等の機会を創出する
0202	障害者の作品制作や表現の技術向上につながる機会が増える 020201 障害者のニーズに合わせ、芸術文化活動を学べる場等、適切な情報提供を行う 020202 障害者福祉施設や学校、地域の文化施設等において、障害者が芸術文化活動に参加・体験できるよう支援する
0203	障害者の芸術文化活動に対する社会的認知が高まる 020301 展覧会や舞台公演等を積極的に開催している 020302 障害者の作品をレンタル、あるいは販売する 020303 多様なメディアを活用し、障害者の芸術文化活動を紹介する
03	障害者の芸術文化活動に従事する人・団体（関係者）がより質の高い芸術文化活動の支援を行うようになる
0301	関係者の必要な知識やスキルが向上し、地域の団体が組織・事業として障害者の芸術文化活動に取り組めるようになる 030101 関係者に障害者の芸術文化活動支援に必要な研修会・勉強会を行う 030102 障害者の芸術文化活動支援に関する成功事例・失敗事例やノウハウ等をテーマ別に蓄積し、発信する 030103 障害者と関係者がともに芸術文化活動の企画を行う
0302	（障害者の芸術文化活動を支援する）多様な主体のネットワークが強化される 030201 様々な機関・専門家が参加した協力委員会等の協働体制を整え、このネットワークを活かした支援を行う 030202 様々なネットワークを活かし、障害者の芸術文化活動支援を行う 030203 障害者の芸術文化活動に関わる行政の担当部署や担当者との定期的な情報交換を行う 030204 必要に応じて、法律専門家や芸術活動専門家の協力を得る
04	地域で障害者の芸術文化活動を応援する人が増える
0401	地域住民が芸術文化活動を通して障害者と交流できるような場がある 040101 地域住民が障害者の芸術文化活動に触れられるような多様な機会（イベント・ワークショップ）をつくる 040102 地域住民が障害者の芸術文化を参加型で一緒に楽しむ機会を創出する 040103 地域住民が障害者の芸術文化を参加型で一緒に楽しむ機会を、地域の団体が創出できるよう支援する
0402	地域住民の障害者の芸術文化活動への関心が高まる 040201 支援センターについて知ってもらうためのリーフレットやウェブサイトを用意し、広報を行う 040202 必要に応じて、障害者の芸術文化活動についてのチラシ等を作成し、地域住民向けに配布する 040203 障害者の芸術文化活動について紹介するウェブサイトを用意し、地域住民向けに公開する 040204 チラシやウェブサイト等の広報媒体は、デザイナーや制作会社に協力を依頼する等して、作成する 040205 地域住民が障害者の芸術文化活動に触れられる多様な機会（イベント・ワークショップ・アーカイブ）についての情報を発信する

(出所)日本ファンドレイジング協会 (2019a) (p.11)

う形で表記する他、「フィデリティ」という用語は使用せず、評価ガイドに示す活動を実施することが、成果達成とどのようにつながるのかということの説明した。また改善のための評価活動を3つのステップに区切り、「ステップ1: 目指すべき事業成果とそこに至る道筋を確認しよう」、「ステップ2: 達成するための道筋に沿って実際に取り組んでみよう」「ステップ3: データをもとに活動状況や成果を振り返ってみよう」として整理した。

「Ⅲ1-3. 効果的プログラムモデルの制度化、制度モデルの作成」と「Ⅲ2-1. 実践現場の創意・工夫と基盤にした効果的プログラムモデルの改訂・更新」に関しては、令和2年度末に評価ガイドを改訂し、「障害者芸術文化活動普及支援事業評価ガイドーより良い協働と事業成果を高めるためのヒント集<sup>8</sup>」として発行された。

## 4. 評価ガイドの活用

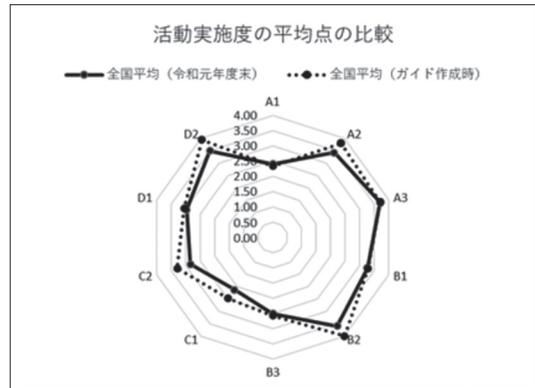
### (1) 活用に向けたはたらきかけ

令和元年度より2年間、開発されたガイドの活用・普及に向けた研修やワークショップの開催を連携事務局や広域センターからの依頼に基づき実施した。令和元年度は、本評価ガイド活用初年度であったため、評価ガイド活用に関する研修を実施し、前提とした評価の考え方等について説明した。

また、研修や成果報告会でフィデリティ尺度<sup>9</sup>を用いた事業の強み、弱みの分析と工夫の共有をするため、評価ガイドに含まれる「支援センター活動のコツ（効果的援助要素）チェックリスト」を各支援センターが活用し、巻末の「レーティング（評点）基準」に沿って活動の実施度（フィデリティ尺度）を測定したものを提出してもらった。そして筆者がその全国の平均値を算出し、評価ガイド作成時と、令和元年度末の情報を比較し、研修等で提示した。その結果を図2に示す。

図2に示されているA1、A2、A3という分類は表3のロジックモデルの4桁（0101、0102、0103）のアウトカムを指しており、Aは4桁の上2桁が

図2 支援センターのフィデリティ尺度



(出所)筆者作成

01、Bは02、Cは03、Dは04を示している。評価ガイド作成時よりも、実施度の平均値が下がっている分野も見受けられたが、大方全国的に同じ傾向であった。研修では、図2で実施度の低いA1（表3のロジックモデルの「0101：地域の障害者に関する芸術文化活動の現状が正しく理解される」にあたる領域）とC1（表3のロジックモデルの「0301：関係者の必要な知識やスキルが向上し、地域の団体が組織・事業として障害者の芸術文化活動に取り組めるようになる」にあたる領域）において、どのような活動を行ったり、課題を感じたりしているかについて支援センター間で情報交換をし、学び合いを実施した。

評価ガイドが導入される以前の成果報告は、各支援センターが独自の視点で「実施した活動内容の報告」を行っていた印象が強かった。しかし、アウトカムを基軸に活動の振り返りをする中で、情報交換や学びの焦点化が起こったように見受けられた。例えば、A1にあたるアウトカム「0101：地域の障害者に関する芸術文化活動の現状が正しく理解される」ためには、支援センター自らが実態調査を仕掛ける方法もあるが、すでにデータを持っている行政との協働事例を他の支援センターから学び、より効果的・効率的なデータ収集ができることを学ぶケースがあった。また、他の支援センターが自分たちとは違う対象に調査をしていることに気づき、限られた資源の中で、どのような人々を対象に調査をすることが、自分たちにとって意義があ

るのかを考えるきっかけを得ている支援センターもあった。

### (2) 活用の現状

令和2年度末に連携事務局が実施した「支援ガイド活用に関するアンケート調査」では、6団体が「活用した」、15団体が「年度末の振り返りで活用する予定」、11団体が「今年度は活用していない」との回答だった。支援ガイドをどの程度活用しているのかについては、「活動のコツのチェックリストを使い、事業の振り返りや事業計画、事業実施に活かした（活かす予定）」が16団体、「ロジックモデルを活用し、事業計画を立てた」が4団体、「評価指標を立てて、事業の振り返りを行った」が6団体であった。

### (3) 支援センターのフィードバック

評価ガイド開発後、2年間の活用において、支援センターの評価ガイドに対する肯定的・否定的意見の双方が聞かれた。肯定的な意見としては、ロジックモデル（インパクトセオリー）を用いて整理することの有用性、活動のコツチェックリスト（効果的援助要素）を活用することによって、他地域と比較した強み・弱みを理解し、事業計画策定に活用できた点、協働型評価によって多角的な視点が導入されたことにより学びが得られた点、活用団体の取り組みが紹介されたことで活用の具体的イメージが持てた点等が指摘された。

否定的意見については、ロジックモデルの作成の困難さや、意義等評価ガイドのデザインに関すること、活動のコツチェックリスト（効果的援助要素）の活用の仕方が目立った。具体的には、「活動のリストの項目の中に複数挙げられている支援対象や施設の「すべて」において活動を行っていないため、項目にチェックをつけていいのかわからない。つけないと「過少評価」になるような気がして、活用が難しい。」というような評価尺度の適切性に関することや、「現場のニーズに合わせて活動しているとリストにない活動も沢山出てくる、それがどのように評価として位置づけるかわからない。」といったリスト外の活動の評価に関することが目立った。

## 5. 評価ガイドの改訂

支援センターにおいて評価が特別なものとして位置するのではなく、より事業改善に役立つものとして事業計画に取り込みやすくしていくために、2年間の活用において見えてきた課題等を整理し、令和2年度末に評価ガイドの改訂を実施した。

改訂における大きなポイントとしては、評価のステップをより事業の運営に内在化させるために組み換えたことや、成果や指標、事業・活動が整理できるワークシートの添付、具体的な指標例の追加がある。これまでの実装の中で、評価ガイドに基づく振り返り以外にも、支援センターが自治体から求められる書類、普及支援事業の年度総括として連携事務局から求められる報告書等、多数の様式が並行していることが明らかとなった。そのため、評価ガイドを活用し事業計画や評価を行うことが、他の書類作成等にもどのように活用できるかについて、改訂の折に支援センターへ付言した。

一方で、全国規模でのロジックモデルや効果的援助要素をアップデートするための機会の創出が叶わなかったため、これらの作業は見送られている。

また2年の間に、評価ガイドのロジックモデル、ないしは効果的援助要素をフォローする支援センター以外に、支援センター独自のロジックモデルを作成し、指標を設定したうえで事業改善に役立てるケースも生まれてきた。このような動きは、CD-TEP法に基づく効果モデルの普及という観点からは逆行するかもしれないが、形成的評価の考え方を現場に浸透させるという意味では、その必要性が理解された結果だと判断した。そしてより支援センターの実態を考慮した運営と、それに基づく計画・改善に役立つロジックモデルを各支援センターが作成する動きは歓迎し、評価ガイドの改訂方針として排除しなかった。

一方、そのような取り組みが増えることで、CD-TEP法に基づく評価ガイドとして、より効果モデルの準拠度は測りにくくなるという問題が発生した。この点に関しては、支援センター全

体に共通するロジックモデルや効果的援助要素に基づいて活動をする団体間で継続的に比較検討に役立てていただくこととした。また、仮に各支援センターが独自でロジックモデルを作成した場合でも、現状では大幅に理念や目的が異なることは想定されにくい（統計的に効果は検証されていないが）参考になる活動は参照していただき、このような多様なアプローチの総体として今後さらに評価モデルの発展を検討していくこととした。結果的に評価ガイドとしては、効果的援助要素は「フィデリティ尺度」としての役割を薄めていくという結果となった。

## 6. 実践の考察

### (1) 形成的評価の促進・定着に関する考察

本実践においてCD-TEP法に基づく評価ガイドが形成的評価の現場の定着に寄与したかどうかに関しては、大きく2つの点からその効果があったと推測される。

1点目は、今回の事例を通じて、現場で行われている実践から事業全体のプログラムセオリーを構築し、支援センターの活動と成果の関係を分析し、効果モデルを生み出したことである。そのモデルをもとに、評価ガイドを作成し、研修等を通じた活用に向けた働きかけを行うことで、成果に基づいて活動を振り返る機会が提供された。このような取り組みを通じて、ただ単に研修の時だけ振り返りを行うのではなく、「活用の現状」にもあるように、実際評価ガイドを活用して振り返りや、事業計画に役立てる支援センターが出現するようになった。「活動のコツのチェックリストを使い、事業の振り返りや事業計画、事業実施に活かした（活かす予定）」の回答者が一番多かったのは、評価ガイドに示す活動ができたか、できなかったのか、したのか、しなかったのかという振り返りは、ロジックモデルを参照・作成したり、評価指標を検討したりするよりも、実施のハードルが低かったからだと推察する。

2点目は、支援センター独自のロジックモデルを作成する団体が出現したことである。評価ガ

イドの活用については、強制的なものではなく、各支援センターに推奨されていたものであったにも関わらず、その意義を感じ日々の運営に取り込んだ団体が現れたことは、ひとつ大きな成果だといえるだろう。

一方、効果的援助要素については、「活動のコツ」リストとして活用している団体もいたが、そのリストにない活動をする団体も確認された。そもそも個別の支援センターが個々のロジックモデルを持てば、アウトカムも変わってくるため、それに基づく活動も異なる。そうすることで、評価ガイド・モデルのフィデリティ尺度としての役割が薄まっていった。この点については、CD-TEP法はモデルの開発を範疇に入れており、効果モデルが広まっていくことへの期待もあったが、支援センターの活動には文脈性があり、本実践の期間中は、それを超えて活動や成果を抽象化し、フィデリティ尺度やロジックモデルのブラッシュアップをすることが出来なかった。

その要因として考慮されるもののひとつは、評価ガイドを発展させていく協働のガバナンス体制構築の難しさである。支援センターは自治体からの委託や補助、広域センターや連携事務局は厚生労働省からの補助を受けて事業を行うため、採択団体の流動性が高い。CD-TEP法自体はボトムアップ評価（Chen 2010）であり、現場の実践からの積み上げを重視する。一方で、普及支援事業の体制には、厚生労働省一自治体一支援センターという典型的な行政の委託・補助事業としての構造と、厚生労働省一連携事務局一広域センター一支援センターという複雑な支援構造が併存しており、限られた事業期間や資源の中で、ボトムアップ評価に基づき、評価ガイドを発展させていく機構の構築が未発達であったといえるだろう。

また、仮に評価ガイドを発展させていく体制が整ったとしても、フィデリティ尺度のアップデートやロジックモデルの改訂といった評価の専門的知見を内製化していくことは、過去3年間の取り組みにおいても外部の評価専門家が求められていたことから、簡単なことではないだろう。

源ほか (2020:27)は、「プログラムの「介入方法」の妥当性について、現実の社会の中で実施しながら検証を繰り返す必要がある」としている。評価ガイドの開発と活用のプロセスを振り返ると、形式的評価の定着については一定寄与したといえるが、「実施しながらの検証」は十分でなかったといえる。しかしながら、このような結果が、本事例において特殊なものであったのか、CD-TEP法という評価アプローチの特有の課題であったのかについては、これまでの考察では導き出せない。一方で、本事例が依拠したCD-TEP法はプログラム理論 (T) の評価と再構築、科学的根拠 (E) に基づく知識生成、実践現場 (P) の創意工夫・改善点の反映を通じて、効果的プログラムモデルの構築を行い、このサイクルを回し続けることで、最終効果モデルを作るもの (大島ほか 2019) としている。本評価ガイドの活用の目指す先が、CD-TEP法が重視する形式的評価の活用と効果モデルの全国的な普及だとしたときに、本評価ガイドの全面的な実装は果たして可能なのであろうか。この点について検討するため、本事例について実装科学の観点から考察し、CD-TEP法の実装可能性について検討する。

## (2) 実装可能性に関する考察

CFIRはDamschroderほか (2009=2021) によって発表され、心理学、社会学、組織変化論等の19の理論やモデルを統合したメタ理論的なフレームワークであり、その源流はRogers (1962) の「イノベーション普及理論」とされている。保健・医療・福祉分野において実装研究を推進するうえで考慮すべき観点が網羅されており、特に実装の阻害・促進要因を特定するための強力なツールとなり、「介入の特性」「外的セッティング」「内的セッティング」「個人特性」「プロセス」という5領域39の構成概念から成り立っている (Damschroder et al. 2009=2021:2)。

今回、筆者は普及支援事業の運営主体や評価ガイドを活用する当事者ではないという点で、すべての概念からの考察は困難であるが、評価デザインや活用推進に携わった立場で一定考察が可能であると推測される「I. 介入の特性」の構成概念に基づいて、CD-TEP法を用いた評価ガ

イドによる介入を考察し、実装にあたっての今後の課題を整理する。「介入の特性」とは、「実装 (実践) の成功の可否に影響する介入の特性のこと」 (Greenhalgh et al. 2004=2021:19, Rabin et al. 2008=2021:19) を指し、「介入の出処」「エビデンスの強さと質」「相対的優位性」「適応性」「試験可能性」「複雑性」「デザインの質とパッケージ」「費用」の8つの構成概念から成る。

### ① I-A 介入の出処

CFIRでは「介入には、良い思いつき、問題の解決、その他の草の根の取り組みから自発的に発生する場合と、外部組織 (業者や研究グループ) によって開発される場合がある (Greenhalgh et al. 2004=2021:20)」とされている。そして、「介入の開発が外発的なものか自発的なものかに関する主要なステークホルダーの認識」については、外発的に開発された介入は意思決定のプロセスの不透明さと相まって実装の失敗につながる場合があるとしている。一方で、「基本構想が外部で開発されても、組織にうまく合わせられれば実装の成功をもたらす可能性がある」 (Gustafson et al. 2003=2021:20) とされている。

本事例では、実際に主要なステークホルダーの認識を確かめてはいない。そのため、あくまでも外部者による見立てに過ぎないが、3年間の関わりにおいて、評価ガイドの開発当初からヒアリングや意見交換等に関わっていた支援センターや、研修やワークショップに参加することで、「自分たちの評価ガイドやモデルを、自ら作成した」と認識・実感できている支援センターは、評価ガイドの活用・モデルの利用に対して、ポジティブな見解を示していた団体が多い印象であった。CD-TEP法自体がボトムアップ型の円環的対話によってモデル構築を促進している点で、効果的な評価ガイドの開発や運用という「介入」開発においては、活用主体がイニシアチブを取ることが重要であり、普及に向けても重要なポイントであることがわかる。特に評価ガイドの活用支援においては、普及支援事業の実施主体が自治体となり、都道府県において一つの支援センターが長く従事している地域もある一方で、支援センター事業を開始して間もない

センターや、自治体内で他の支援センターが採択されるケース、もとより、自治体の担当者も変化するということが起こっていた。使い手の移ろいは常に新たな視点をもたらす可能性があるが、コミットメントやエンゲージメントを高めていかない限り、常に活用において「お客さん」になってしまう可能性が否めず、CD-TEP法の効果モデルの改善・普及のガバナンスの在り方については今後も検討を深める必要があると推測される。

## ② I-B エビデンスの強さと質

Damschroderほか(2009=2021)は強力なエビデンスは重要であるものの、介入を採用する個々の決定に必ずしも影響力があるわけではないとしている。これは効果モデルを普及していく時に、「これは成果の達成に効果が認められたエビデンスの高い活動だ」ないしは、「この評価ガイドに基づく活動をすることで、成果が高められる」ということをアピールすることは、介入の拡散・普及には寄与しないということを示しているように解釈できる。一方で、Rycroft-Maloneら(2002=2021:21)は「臨床においては、PARIHSモデルは研究知見の採用のカギとなるエビデンスの3つの採用条件として、調査研究、臨床経験、および患者の体験」を挙げており、CFIRではこのようなエビデンスの情報源が多いほど、イノベーションが受け入れられる可能性が高まるとしている。本事例に読み替えると、評価ガイドのデザインに係る調査研究(筆者が実施)、評価ガイドの使用経験と、活用した支援センターの事例等が該当する。強力なエビデンスは重要である一方で、介入を受けた(評価ガイドを活用した)支援センターの使い勝手の向上や、改善プロセスの体験といった情報は、介入の普及という点で必要である可能性が高く、CD-TEP法を用いて開発した評価モデルを実装していく際に参考になる可能性がある。

## ③ I-C 相対的優位性

Gustafsonほか(2003=2021:22)は、相対的優位性を、「代わりの解決策と比較した、介入を実装することの強みに関するステークホルダーの認

識」としており、ステークホルダーが介入の有効性や効率に明らかな優位性を認識していれば、実装は成功する可能性が高いとしている(Damschroder et al. 2009=2021)。また、Damschroderほか(2009=2021:22)は多くの研究に基づき、介入の有益性は、相対的優位性を評価するためにステークホルダーに明確に見えるものでなければならず、介入の有益性を明示するための努力が、実装を促進するために必要であるとしている。本事例においては、例えば評価ガイドにおいて、従来型の評価や協働型の評価ではない評価と本モデルを比較して、本評価ガイドの意義を伝えた他、「使ってみた」「やってみた」事例を紹介・解説することで、「できるかも」「やる意義がありそう」だという認識に少しずつ変えていくプロセスが、特に評価ガイドの活用フェーズに起きていた。「① I-A 介入の出处」で述べたとおり、CD-TEP法は円環型対話でモデルがブラッシュアップされる特性があるが、本事業のように支援センターの経験年数や、コミットメント度合い、自治体の温度感等によって必ずしもすべてのステークホルダーが円環的対話の席に常についているとは限らない。このような観点から、普及に向けた活動には相対的優位性を意識した関わりも一定必要になる可能性があると考えられた。しかし、「実施するお徳感」のようなものが過度に表現されることで誤った方向にインセンティブが働く可能性もあり、努力の度合いは将来的に懸念事項となるだろう。

## ④ I-D 適応性

適応性は現場のニーズを満たすために、介入の適応やカスタマイズ、洗練化、見直しが可能なる程度を指す(Damschroder et al. 2009=2021:23)。CD-TEP法は形成的評価を繰り返しながら、効果モデルを現場の声をもとにアップデートすることが可能である。そのため、介入手法として適応性は高いと考えられた。しかし、評価ガイドが示す効果的援助要素(フィデリティ尺度)をどの程度重要視し、評価や事業計画を策定するかが、実装中に実際の課題として浮き彫りとなった。適応性は介入の「中心的要素」と「適応可能な側面」の定義によって決まるとある

(Greenhalgh et al. 2004=2021:23)。CD-TEP法に基づいた評価ガイドを貫き通すことは当然ながら現場の本義ではないため、どの程度のバランスまでが評価アプローチとして許容されるかという課題が残った。

#### ⑤ I-E 試験可能性

試験可能性は介入を組織で小規模で試験し、妥当な場合はその逆行ができることを指す(Damschroder et al. 2009=2021:24)。CD-TEP法のアプローチとして試験可能性は備えられていると考えられ、実際のプロセスにもトライアルを重ねることが示唆されている。一方で、評価ガイドの作成が直面した課題もあった。それは、今回のフィデリティ尺度を作成したときは、アンケートが戻った41団体を対象に統計処理を行ったが、アウトカムと活動の相関については標本数が少ないと正しい分析ができないという問題である。Greenhalghほか(2004=2021:24)は、「介入の事前テストが行えることと、効果的な実装は強い正の関連があります」としているが、このような結果に基づいて開発された介入については常に考慮が必要である。

#### ⑥ I-F 複雑性

複雑性は実装の困難さについての認識であり、具体的には実装に必要な複雑さと手順数等を指す(Damschroder et al. 2009=2021:25)。CFIRでは「急進的な介入は、組織活動に根本的な変化が生じる大幅な方向転換や非日常的プロセスが必要であり、従来の慣習からの明確な脱却をもたらす」(Greenhalgh et al. 2004=2021:25)とあり、評価ガイドの活用は、このようなドラスティックな変化自体を目指したものではない。しかし、今回の事例においては、①結果の評価から成果の評価へ、②事業計画における評価視点の導入、③ロジックモデルの作成(必須ではない)、④評価のための指標の選定、⑤活動のコツチェックリスト(効果的援助要素)のチェック、⑥これらを議論するための追加マネジメント等が結果として必要となっている。

社会的インパクト・マネジメント<sup>10</sup>の観点からは、今回のような評価のプロセスは、事業マネ

ジメントのプロセスに併合できるという考えもある。一方で、従来あった「評価への抵抗感」を考慮すると、いくら事業を改善するものとしても、追加の工数を求められているという認識はぬぐえない可能性があった。そのため、活用促進の観点からは、最初からフルパッケージでやるのではなく、できるところからやってみるという形で進めた。普及にあたっては介入を受ける(評価を活用する)側のレディネスのアセスメント等も必要になるだろう。

#### ⑦ I-G デザインの質とパッケージング

これは介入がうまくまとめられており、提示、組み立てが優れていることに関する認識のことである(Damschroder et al. 2009=2021:27)。CD-TEP法においては、サービス利用計画やツールキットを作成することが明示されており、使い手を意識したデザインとなっている。本事例においても、「評価ガイド」という形で提示し、特に読み手を意識した内容とした。また実用性を重視し、改訂を行っている。

#### ⑧ I-H 費用

介入の費用、および介入実装に関連する費用は実装と負の関係にあるとされている(Graham and Logan 2004=2021:28)。本事例においては、支援センター自体に原則として金銭的な支出を求めるものではなかったが、支援センターからのフィードバックには「予算(人件費)がないため、出来る業務に限られている」、「ワークショップには参加したが、評価を用いて事業を組み立てなおす時間をセンター内で取ることが出来なかった」という意見が聞かれた。評価の領域では、よく評価費用が問題視されるが、自身の事業を計画する際に、一般的にはその計画のためのコストは事業関連の費用の中に吸収されると考えられるだろう。形成的評価は、特に事業設計の段階から組み込めるものであるため、「評価」としての追加コストと考慮するより、事業費のより効果的な活用方法として、事業費に内在化できる可能性を大いに持っている。しかし、そのように認識しない支援センターが多いことも事実であり、評価導入の複雑さや活用のため

の研修等に拘束されるコスト意識が、普及の阻害因子になる可能性があることは興味深かった。

## 7. CD-TEP法を活用した形成的評価の促進に向けた展望と課題の整理

今回、評価ガイドの開発で採用したCD-TEP法は、形成的評価を効果モデル生成の手段として位置づけていた。事例においては、評価ガイドの開発・活用・改訂プロセスの考察でも述べたように、自らの事業を振り返り、振り返りから出てきた情報をもとに改善を行い、それを次の事業計画等につなげ、その中に評価（ないしは形成的評価）の視点が生きるということを実感し始めた団体が見受けられた。このことから、形成的評価の定着という観点からは一定の成果があったと考えられる。特に、多様なレベルで中間支援を実施している団体によって複雑に構成されている普及支援事業においても、研修等の働きかけにより形成的評価の重要性や意義が伝わっていったことは、他の特徴を共有する事業等においても、形成的評価定着の望みを感じられる結果といえるだろう。

一方で、評価ガイドの活用のプロセスにおいて、結果として「形成的評価における改善」に比重が置かれていったことにより、「効果モデル構築」というCD-TEP法の目的が薄まっていったことは看過できないポイントである。「I. 介入の特性」の「I-D 適応性」でも示した通り、CD-TEP法の効果モデルは可変であり、現場との対話や実践の振り返りにより、モデルをアップデートできる。その理念は理解できるが、効果モデルを展開・普及していくという長い道程を目指していくには、当然ながら理念だけでは不十分であり、それらを前に進めていくための効果的な実装の仕組みや仕掛けが必要であることがわかった。

CFIRの概念には、今回分析に活用した概念以外に、外的セッティング、内的セッティング、個人特性、プロセスといった概念も存在する。外的セッティングは介入を実施する組織がおかれた経済・社会・政治的文脈を指し、内的セッ

ティングは介入を実施する組織内部の構造的、政治的、文化的文脈を示している (Damschroder et al. 2009=2021:29-59)。また、個人的特性は介入の対象となる人々の特性を指し、プロセスは介入が個人のレベルと組織のレベルで計画通りに行われるようにするための積極的な活動プロセスを指している (Damschroder et al. 2009=2021:61-78)。本事業に部分的に関わった筆者ではそれらすべてをレビューすることはできないが、本稿にも示した普及支援事業の体制や構造、運営形態といった条件や、リーダーシップ、個々人の評価に関する認識やスキル等も、考慮すべき事項であることがうかがえる。

このような点から、CD-TEP法が重視するボトムアップの視点は、理念的には尊くとも、それを実現し、実装していくためのガバナンス・体制がなければ、本来の目的を達成しない可能性がある。本事例においては、評価ガイドの開発に着手した際に、「評価ガイドの開発」だけをボトムアップとして考えるのではなく、実装を見据え、ガイド開発のプロセスにもその先の普及においてキーとなる広域センターや連携事務局に関わってもらうことで、実装の展開が変わったかもしれない。このようなプロセスを経て、広域センターと支援センターの「評価」に関するコミュニケーションの在り様も変化した可能性があっただろう。もとより、このような関わりは、相応の手間や時間を要するもので、これら自体をそもそもの役割や職掌等に事前に組み込んでいく視点も重要だろう。

また、本事例では筆者が評価ガイドの開発や活用支援を現場とともに行ったが、実装を支える専門的な関わりなしにはモデル構築は難しいだろう。普及支援事業の実施者の中に、評価の専門家がいることで、この課題は解決される可能性はあるが、裏を返せば、CD-TEP法自体は評価の専門性がある程度なければ、実装自体が難しい評価アプローチだともいえる。しかし、本事例でも幾分明らかになったように、研修等の働きかけにより、現場団体にも形成的評価の意義や方法を伝えていくことで、現場の評価キャパシティを向上させていくことは可能であり、普及支援事業の実施主体が自らの実践を振り返

られる評価の力を身につけながら、事業全体としての評価キャパシティ形成が促進されることが望まれる。本事例においては、支援センターをサポートする広域センターや連携事務局がその役割を果たせる可能性がある。このような評価ガイドや効果モデルづくりは相応のエンゲージメントが求められることもあり、実践現場における評価ガイドの開発は始まりに過ぎず、実装支援の仕組みや体制整備と相まって展開される必要があるだろう。

## 注記

- 1 令和4年4月1日時点での所属であり、評価ガイドの開発と研修等は筆者が特定非営利活動法人日本フアンダレイジング協会所属時に受託したものである。
- 2 一般的な略称は「CD-TEP評価アプローチ法」であるが、本稿では「CD-TEP法」と記した。
- 3 <https://www.mhlw.go.jp/content/12200000/000521749.pdf> (2022年6月12日閲覧)
- 4 中間支援組織の定義は多様であるが、平成23(2011)年の内閣府の『新しい公共支援事業の実施に関するガイドライン』では、「市民、NPO、企業、行政等の間にたって様々な活動を支援する組織であり、市民等の主体で設立され、NPO等へのコンサルテーションや情報提供などの支援や資源の仲介、政策提言等を行う組織を言う。中間支援組織自らがNPO等である場合もある」と定義している。
- 5 [https://www.radish-japan.org/files/CFIR\\_Guidebook2021.pdf](https://www.radish-japan.org/files/CFIR_Guidebook2021.pdf) (2022年6月12日閲覧)
- 6 社会的プログラムの効果的な実施に関わるプログラムの実施内容・実施要素を具体的に示したものであり、主要なアウトカム改善に寄与するプログラム要素はエビデンスに基づく実証研究によって明らかにされるとともに、効果モデルのプログラム理論、インパクト理論、プロセス理論にも基づいて構築される(大島ほか2019)。
- 7 中の「○」は実施した項目、「△」は一部実施した項目、「－」は実施していない項目を示している。
- 8 <https://arts.mhlw.go.jp/wp-content/uploads/2021/07/c3f4d37f7aefcc2046f3da4aa5ea1929.pdf> (2022年6月12日閲覧)
- 9 評価ガイドではCD-TEP法に基づき、ロジックモデ

ルのアウトカムとつながっている活動が、本当にその達成に寄与しているのかについて、アンケート調査のデータをもとに検証している。その結果、おおよそロジックモデルに示した仮説の通り、意図したアウトカムとそれを達成するための活動は互いに関連しあっており、効果的援助要素(ロジックモデルに示す活動)の実施度(フィデリティ尺度)が目標とする4点(大島2012)を達成している団体は、4点未達成の団体に比べて、より良いアウトカムを達成している傾向にあることが確認されている。そのため、支援センターにおける障害者芸術文化活動支援を効果的に実施するためには、効果的援助要素にある様々な活動に取り組むことが、アウトカムの達成に寄与することが有効であるというロジックが成り立っている。

- 10 社会的インパクト・マネジメント・イニシアチブ(2021)では社会的インパクト・マネジメントを「事業や取り組みがもたらす変化や価値に関する情報を、各種の意思決定や改善に継続的に活用することにより、社会的インパクトの向上を目指す体系的な活動のこと」と定義している。

## 参考文献

- 大島巖(2011)『CD-TEP円環的対話型評価アプローチ法実施ガイド』、<http://cd-tep.com/> (2022年3月31日閲覧)
- 大島巖(2012)『CD-TEP円環的対話型評価アプローチ法実施ガイド 平成22年度文部科学省・科学研究費補助金基盤研究(A)「プログラム評価理論・方法論を用いた効果的な福祉実践モデル構築へのアプローチ法開発」報告書』
- 大島巖・源由理子・山野則子・他(2019)『実践家参画型エンパワメント評価の理論と方法—CD-TEP法:協働によるEBP効果モデルの構築』、日本評論社
- 厚生労働省(2021)『障害者芸術文化活動普及支援事業』、厚生労働省
- 社会的インパクト・マネジメント・イニシアチブ(2021)『社会的インパクト・マネジメント・ガイドライン VERSION 2』、社会的インパクト・マネジメント・イニシアチブ
- <https://simi.or.jp/wp-content/uploads/2021/07/Social-Impact-Management-Guideline-ver.2.pdf> (2022年6月

- 12日閲覧)
- 内閣府 (2011)『新しい公共支援事業の実施に関するガイドライン』、内閣府
- 日本ファンドレイジング協会 (2019a)『障害者芸術文化活動普及支援ガイドーより良い協働と事業成果を高めるためのヒント集一』、日本ファンドレイジング協会
- 日本ファンドレイジング協会 (2019b)『障害者芸術文化活動普及支援事業評価ガイドライン作成事業報告書』、日本ファンドレイジング協会
- 日本ファンドレイジング協会 (2021)『障害者芸術文化活動普及支援事業評価ガイドーより良い協働と事業成果を高めるためのヒント集一』、日本ファンドレイジング協会
- 源由理子・大島巖 (2020)『プログラム評価ハンドブックー社会課題解決に向けた評価方法の基礎・応用』、晃洋書房
- 令和2年度障害者芸術文化活動普及支援事業連携事務局 (2021)『令和2年度障害者芸術文化活動普及支援事業報告書』、令和2年度障害者芸術文化活動普及支援事業連携事務局
- Chen, H.T. (2010). The bottom-up approach to integrative validity: A new perspective for program evaluation. *Evaluation and Program Planning*, 33 (3), 205-214.
- Damschroder, L. J., Aron, D. C., Keith, R.E., Kirsh, S.R., Alexander, J. A., and Lowery, J.C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, 4 (1), 50-65. (2021、今村晴彦、島津太一 (訳)『実装研究のための統合フレームワークーCFIRーConsolidated Framework for Implementation Research』、保健医療福祉における普及と実装科学研究会)
- Graham, I. D., and Logan, J. (2004). Innovations in knowledge transfer and continuity of care. *Canadian Journal of Nursing Research*, 36 (2),89-103. (2021、今村晴彦、島津太一 (訳)『実装研究のための統合フレームワークーCFIRーConsolidated Framework for Implementation Research』、保健医療福祉における普及と実装科学研究会)
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., and Kyriakidou, O. (2004). Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Quarterly*, 82 (4), 581-629. (2021、今村晴彦、島津太一 (訳)『実装研究のための統合フレームワークーCFIRーConsolidated Framework for Implementation Research』、保健医療福祉における普及と実装科学研究会)
- Gustafson, D. H., Sainfort, F., Eichler, M., Adams, L., Bisognano, M., and Steudel, H. (2003). Developing and testing a model to predict outcomes of organizational change. *Health Service Research*, 38 (2), 751-776. (2021、今村晴彦、島津太一 (訳)『実装研究のための統合フレームワークーCFIRーConsolidated Framework for Implementation Research』、保健医療福祉における普及と実装科学研究会)
- Rabin, B. A., Brownson, R. C., Harie-Joshu, D., Kreuter, M. W., and Weaver, N.L. (2008). A Glossary for Dissemination and Implementation Research in Health. *Journal of Public Health Management and Practice*, 14 (2), 117-123. (2021、今村晴彦、島津太一 (訳)『実装研究のための統合フレームワークーCFIRーConsolidated Framework for Implementation Research』、保健医療福祉における普及と実装科学研究会)
- Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press of Glencoe.
- Rycroft-Malone, J., Harvey, G., Kitson, A., McCormack, B., Seers, K., and Titchen, A. (2002). Getting evidence into practice: Ingredients for change. *Nursing Standard*, 16 (37),38-43. (2021、今村晴彦、島津太一 (訳)『実装研究のための統合フレームワークーCFIRーConsolidated Framework for Implementation Research』、保健医療福祉における普及と実装科学研究会)

(2022.6.30 受理)

## **Development and Implementation of an Evaluation Guide for Developing Field Practices into an Effective Model － A Case Study Using the CD-TEP Approach －**

Junko Shimizu

Musashino University  
jshimizu@musashino-u.ac.jp

Kenta Shindo

Japan College of Social Work  
shindo@jcs.w.ac.jp

### **Abstract**

This article examines a case study of the development and implementation of an evaluation guide for the nationwide intermediate welfare and art support centers for the disabled, which was developed using an approach of the Circular Dialogue between Program, Theory, Evidence, and Practices (CD-TEP). By reviewing that process, this study analyzes how the CD-TEP-based evaluation guide promoted formative evaluation in the field and to what level the approach is implementable. The result shows that the guide utilization could have promoted the formative evaluation practices. However, it was observed that the role of the model effectiveness and fidelity scale in the evaluation guide has diminished. Also, the study on the implementability of the CD-TEP method partially using the framework of implementation science (CFIR) showed that for the philosophy as an evaluation approach to penetrate practice, appropriate governance and systems are needed, and the necessity of technical implementation support was affirmed.

### **Keywords**

Formative Evaluation, CD-TEP, Evaluation Guide, Intermediary, Implementation Science



## 【実践・調査報告】

## 実践家参画型エンパワメント評価の基盤を支える 「EBP 効果モデル」技術支援センターの意義と役割 — 評価キャパシティ形成に向けた役割を中心に —

新藤 健太

日本社会事業大学  
shindo@jcs.ac.jp

大島 巖

東北福祉大学  
i-oshima@nifty.com

### 要 約

効果的・普遍的なプログラムモデルの構築とその実施・普及に有効なプログラム評価・開発のアプローチ法として実践家参画型エンパワメント評価がある。この取組みの推進には実践家及び組織の評価キャパシティ形成が欠かせない。そこで本稿では2つの事例をもとに、評価キャパシティ形成を支える「EBP効果モデル」技術支援センターの意義と役割を検討した。

その結果、評価キャパシティ形成のアウトカムには実践家及び組織レベルの変化があること、その先には、サービス・アウトカムや利用者アウトカムの向上もあること、そのために「EBP効果モデル」技術支援センターは、①協働による効果モデル開発の取組み、②アセスメントと評価キャパシティ形成のための計画策定、③具体的な支援メニュー（トレーニング・コンサルテーション・見学等の機会）提供、④ネットワークを行うことが有効であると考察された。

### キーワード

EBP効果モデル技術支援センター、評価キャパシティ形成、実践家参画型エンパワメント評価

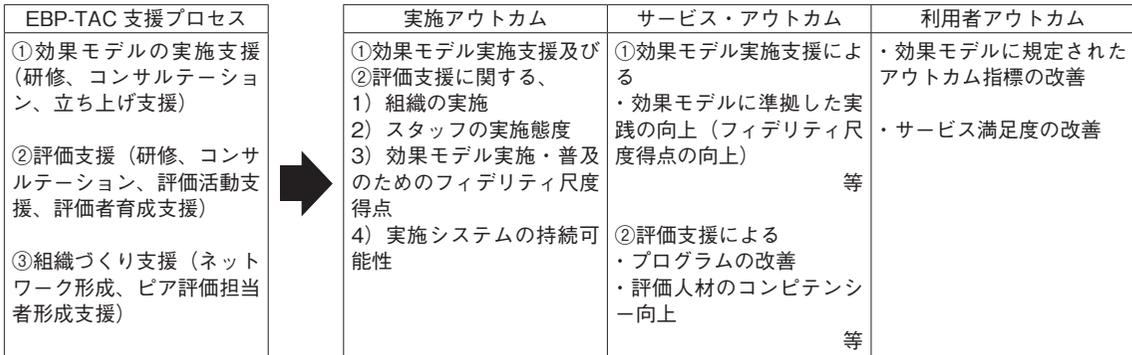
#### 1. はじめに

近年、形成的評価やプログラム開発に対する関心が高まっている。これらに取組むうえで求められるのは、新しく効果的なプログラムを開発し、より有効なものへと改善するための理論と方法論であり、実践家参画型エンパワメント評価（大島ら 2019、pp.1-14）はその一つであるといえる。実践家参画型エンパワメント評価とは、大島ら（2019）が開発した「EBP（Evidence-Based Practices）等効果モデル」構築のための方法論で、実践と科学的なエビデンスに基づいて、

まずは実践家（プログラムの提供者）がプログラム評価・開発活動に積極的に参画し、さらには利用者（プログラムの受益者）や市民、その他関係者たちの協働を得て、一緒に取組みを進めることに特徴がある（大島ら 2019、p.1）。これまで、大島ら（2019）の研究チームはこの方法論を用いて、いくつかの普遍的・効果的なプログラムモデル（効果モデル）の開発に取組み、成果を上げている。

このような実践家参画型エンパワメント評価の特徴から、実践家等の評価キャパシティ形成（ECB：Evaluation Capacity Building）に向けた取

図1 EBP-TACが行う評価キャパシティ形成 (ECB) のロジックモデル



(出所) 大島ら (2019, p.255) をもとに筆者作成

組みが重要になる。ただし、評価キャパシティ形成 (ECB) は多くの場合容易ではなく、実践家等のみで誰の手助けも得ずに取組みを進めていくことは相当な困難があるものと考えられる。そのため、このような役割を担う機関・組織として「EBP効果モデル」技術支援センター (EBP-TAC: Evidence-Based Practices Technical Assistance Center) が位置付けられる。しかし、現状、我が国においては、評価キャパシティ形成 (ECB) の具体的方法論についても、「EBP効果モデル」技術支援センター (EBP-TAC) が担うべき役割の整理とその実装・設置についても十分な研究的・実践的な経験の蓄積はない。そこで本稿では、我が国における2つの事例を踏まえながら評価キャパシティ形成 (ECB) に果たす「EBP効果モデル」技術支援センター (EBP-TAC) の役割について検討する。

## 2. 本研究の視点と分析方法

### (1) 評価キャパシティ形成 (ECB)

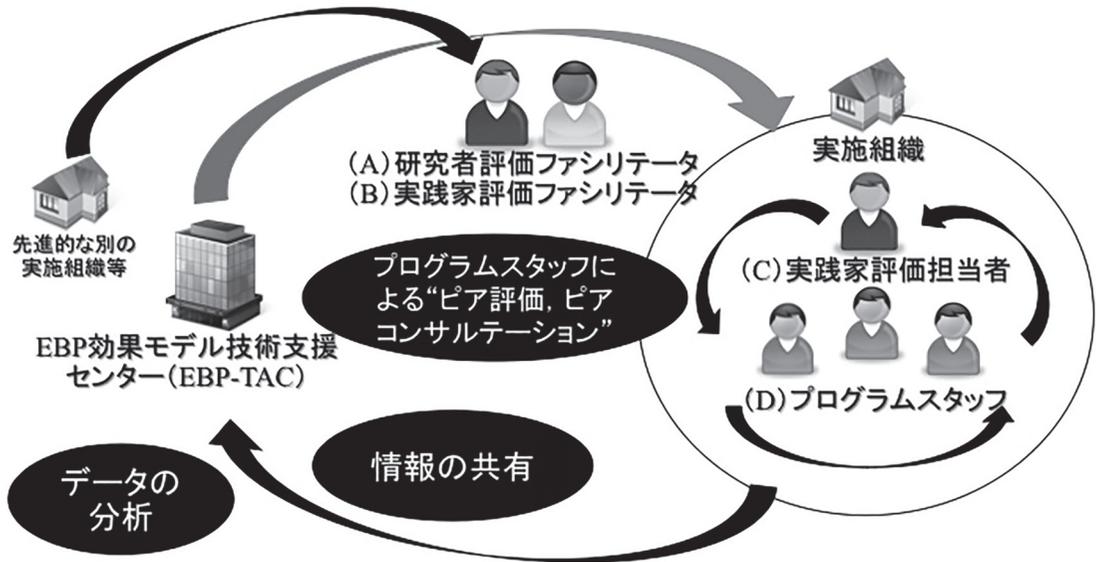
評価キャパシティ形成 (ECB) には「個人レベル」、「組織レベル」のものがある (Preskill and Boyle 2008; Labin et al. 2012; 大島ら 2019)。具体的に、Preskill and Boyle (2008) は個人レベルの評価キャパシティ形成 (ECB) の目的、すなわち ECBによって得る個人レベルのアウトカムを、知識 (評価に関するいくつかの事項を知っていること)、スキル (評価に関するいくつかの事項を

実践できること)、態度 (評価に関わるいくつかの事項を信じられること) の3つから整理している。

また、Labin et al. (2012) は組織レベルの評価キャパシティ形成 (ECB) の目的を、プロセス・ポリシー・プラクティス (3つのP) (組織における「評価の増加」や「評価の活用機会増加」等)、リーダーシップ (評価に関するサポートの増加)、組織文化 (組織文化として「学習・参加・協働」や「評価・データの活用」があること)、主流化 (「評価の日常化」や「業務への評価の取り込み」、等を実現すること)、資源 (「評価に関する資料」や「テクノロジー」、「組織内部の評価に関する専門知識」等を得ること) の5つから整理している。

これらの定義を踏まえて大島ら (2019, pp.56-57) は、実践家参画型エンパワメント評価における評価キャパシティ (EC) を「プログラムに関わる実践家やその他関係者が評価を体系的に実施し、それとともに評価を活用してプログラムの計画を立案し、プログラムを改善する個人及び所属する組織の能力」とし、評価キャパシティ形成 (ECB) を「このような能力を高める意図的なプロセス」と定義している。さらに大島ら (2019, p.255) は、Proctor et al. (2009) を参考に、「EBP効果モデル」技術支援センター (EBP-TAC) のロジックモデルを整理し、評価キャパシティ形成 (ECB) に関連したアウトカムを「実施アウトカム」、「サービス・アウトカム」、「利用者アウトカム」の3段階から捉えている (図1)。

図2 「EBP効果モデル」技術支援センター（EBP-TAC）の概要



(出所) 効果のあがる就労移行支援プログラムのあり方研究会（2015、pp.11-12）をもとに筆者作成

このロジックモデルは2段階目の「サービス・アウトカム」に効果モデルの実施度向上（フィデリティ尺度得点の向上）やプログラムモデルの改善を位置付けている点に特徴がある。ただし、この「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデルについては、現状では十分な実践的裏付けが得られておらず、実践事例との対照を通して、規定された取組みは実施可能か、取組みが実施されれば意図した成果（アウトカム）を得ることができるのかといった検証が必要である。

(2) 「EBP効果モデル」技術支援センター（EBP-TAC）

実践家参画型エンパワメント評価は形成的評価を用いながら、実践家等の関係者と協働してEBP等効果モデルの開発を目指す。そのため、実践家等とのネットワークを構築し、実践家等の評価活動への参画を促すことや、実践家等の評価キャパシティ形成（ECB）等を目的とした中間支援組織の存在が必要になる。実際、海外の先進的な取組みではこのような中間支援組織の存

在がエビデンスに基づく知見づくりや、エビデンスの活用促進、評価キャパシティ形成（ECB）に大きく貢献している実態がある（IES 2011；大島ら 2019）。実践家参画型エンパワメント評価では、こうした海外の先進的事例等も参考にしながら「EBP効果モデル」技術支援センター（EBP-TAC）の仕組みづくりにも取り組んでいる。その概要を図2に示す。「EBP効果モデル」技術支援センター（EBP-TAC）及び実践家参画型エンパワメント評価に必要な人材を、大きく、(A) 研究者評価ファシリテーター、(B) 実践家評価ファシリテーター、(C) 実践家評価担当者、(D) その他のプログラムスタッフの4者から説明する（効果のあがる就労移行支援プログラムのあり方研究会 2015；大島ら 2019）。まず、(A) 研究者評価ファシリテーターとは福祉系大学等、すなわち「EBP効果モデル」技術支援センター（EBP-TAC）に所属し、実践家参画型エンパワメント評価に関する知識と技術を十分に備えた人材である。実践家評価ファシリテーターの活動を後方から支援する役割も担うため「コンサルタント評価ファシリテーター」と呼ぶこともできる。

次に、(B) 実践家評価ファシリテーターとは実践に関わりながら評価ファシリテーターとしての役割を果たす人材である。彼ら・彼女らは、実施組織と関わり、プログラムの実施支援や評価支援を行う場合、実践家評価担当者及びその他のプログラムスタッフとは同じ「実践家」という立場に立って評価ファシリテーターとしての役割を果たすことから「ピア評価ファシリテーター」と呼ぶこともできる。なお、実践家評価ファシリテーターは通常、実施組織にてプログラムの実施に取り組んでいることから日常的に「EBP効果モデル」技術支援センター（EBP-TAC）に所属しているわけではなく、実施組織に対するプログラム実施支援や評価支援等に際して、その都度、「EBP効果モデル」技術支援センター（EBP-TAC）の業務に従事する人材である。次に、(C) 実践家評価担当者とは、プログラムを導入する実施組織でプログラム実施及び評価活動の中心的な役割を担う人材である。研究者・実践家評価ファシリテーターと連携しながら、その他のプログラムスタッフと協働してプログラムの実施及び評価活動を遂行する。最後に、(D) その他のプログラムスタッフは、プログラムを導入する実施組織内で、実践家評価担当者に協力する形でプログラムの実施や評価活動に取り組む人材である。

これらの人材が協働し、評価を活用しながらプログラムの実施に取り組む。そして、評価から得られた様々な情報は「EBP効果モデル」技術支援センター（EBP-TAC）で集約・分析され、プログラムの改善（EBP等効果モデルの構築）に生かされる。そうして改善されたプログラムは、研究者・実践家評価ファシリテーターを介して実施組織へと伝達され、実践家評価担当者を中心に、新たな実践が展開されることで、実施組織における実践も改善し、利用者（プログラム受益者）のアウトカムもより良く向上するようになる。こうした営みを支えるのが「EBP効果モデル」技術支援センター（EBP-TAC）である。なお、本稿においては特に実践家等の評価キャパシティ形成（ECB）に焦点を当てていることから主に、「EBP効果モデル」技術支援センター（EBP-TAC）による（B）実践家評価ファシリテ

ーターや（C）実践家評価担当者、（D）その他のプログラムスタッフを対象にした働きかけと、彼ら・彼女らの変化に焦点を当てる。

### (3) 分析方法

本稿の目的である評価キャパシティ形成（ECB）に果たす「EBP効果モデル」技術支援センター（EBP-TAC）の役割を検討するために、実践家参画型エンパワメント評価による取組みのうち、筆者が関与してきた「効果的障害者就労移行プログラム」及び「生活困窮者自立相談支援事業プログラム」の2事例を分析対象として取り上げ、それぞれに評価キャパシティ形成（ECB）の取組みを分析する。具体的にはそれぞれの事例に対して、実践家参画型エンパワメント評価の取組みを進めていく中で、評価キャパシティ形成（ECB）に関して、誰にどのような働きかけを行ったか【視点①】、その結果評価キャパシティ形成（ECB）はどのように進んだか【視点②】、さらにこれがプログラムモデル（効果モデル）の構築にどう貢献したか【視点③】を明らかにする。

本稿では、これらの視点による分析を通して、大島ら（2019, p.255）が作成した「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデルは実際の事例に当てはまるものなのかを検証する。さらに、この分析を通して「EBP効果モデル」技術支援センター（EBP-TAC）が評価キャパシティ形成（ECB）に果たすべき役割、期待、そして今後の課題について考察することを目的とする。

## 3. 結果とその考察：2事例に対する分析

### (1) CD-TEP法：2事例に用いた効果モデル開発の研究手法論

実践家参画型エンパワメント評価ではEBP等効果モデルの開発とその実施・普及のためのアプローチ法として「プログラム理論・エビデンス・実践間の円環的対話による効果的プログラムモデル形成のためのアプローチ法（CD-TEP）」（大島ら2019）を用いる。

CD-TEP法は、新しく導入される対人サービスの実践プログラム、あるいは十分な成果が上げられていない既存プログラムを、プログラムに関わる実践家や利用者等関係者の参加と協力を得て、より効果的で有用性の高いプログラムモデル（効果モデル）へと発展させるために行う、ボトムアップ型プログラム評価・開発の方法論である（大島ら 2019、p.76）。

なお、実践家参画型エンパワメント評価（CD-TEP法）における「効果モデル」という用語は、インパクト理論、プロセス理論（サービス利用計画・組織計画）、効果的援助要素リスト、評価ツール（フィデリティ尺度、アウトカム測定ツール）、プログラム実施マニュアルといった5つのアイテムを総称したものと用いられる（大島ら 2019、pp.94-95）。

さらにCD-TEP法では、プログラムモデル（効果モデル）構築のための手順を【ステップ1】現状分析・ニーズ評価：問題状況とニーズの分析、ターゲット集団とプログラムゴールの設定、【ステップ2】評価可能性アセスメントの実施と予備的效果モデル（暫定版）の作成、【ステップ3】GP（Good Practice）事例調査の実施、【ステップ4】質的データ分析と実践家等参画型ワークショップの準備、【ステップ5】実践家等参画型ワークショップ：第1次効果モデル（試行版）構築、【ステップ6】第1次効果モデル（試行版）の形成・構築・改善：効果モデル5アイテム（インパクト理論、プロセス理論、効果的援助要素リスト、評価ツール、実施マニュアル）の作成、【ステップ7】広域的事業所調査：第1次効果モデル（試行版）の広域的な検証、【ステップ8】広域的試行評価調査①：単一グループデザインで行う多施設共同調査、【ステップ9】質的・量的データ分析と実践家参画型ワークショップの準備、効果モデル改定案の作成、【ステップ10】実践家等参画型ワークショップ：第2次効果モデル（提示版）への形成・改善、【ステップ11】広域的試行評価調査②：比較による有効性研究（CER）で行う多施設共同調査、【ステップ12】第3次効果モデル（エビデンス版）への形成・改善：質的・量的データ分析と実践家等参画型ワークショップによる効果モデル（エビデンス版）への形成・

改善から整理している（大島ら 2019、pp.131-141）。

## (2) 効果的障害者就労移行支援プログラム

### ①プログラムモデル（効果モデル）構築のプロセス

障害者就労移行支援事業には、事業所全体の3割弱が0%の就労移行率である等、十分な成果を上げていない状況がある（新藤ら 2017、p.57）。そこで筆者らの研究チームはCD-TEP法を用いてこの事業の効果モデル構築に取り組むこととした。

効果モデル構築の取組みは2007年度に開始され、いくつかの先行研究を参考に仮説モデルを構築した。次に、就労移行率等の成果（アウトカム）が高い事業所10か所をGP（Good Practices）事例として抽出し、訪問調査を実施した【ステップ1～3に該当】。その後、このプログラムのインパクト理論やプロセス理論、効果的援助要素リストを検討するためのフォーカス・グループインタビューを行い、第1次効果モデル（試行版）を構築した【ステップ4～6に該当】。2009年には、全国1,213か所の就労移行支援事業所を対象に、第1次効果モデル（試行版）に基づく広域的事業所調査（アンケート調査）を実施し、フィデリティ尺度得点がアウトカム指標と正の相関関係にあることを確認した【ステップ7・9に該当】。その後、2009年2月から2009年12月にかけて、2013年11月から2014年11月にかけての期間、全国の就労移行支援事業所数十か所を対象にした全国試行評価調査①を実施した。この間、数回にわたって実践家参画型ワークショップも開催し、効果モデル改訂に向けた検討を実施した。この調査から得たデータ分析の結果、フィデリティ尺度得点の高いグループ（効果的援助要素の実施度が高いグループ）の方が、これが低いグループに比べてより良いアウトカムの値を示していることも確認した（新藤ら 2017）【ステップ8・9に該当】。これらの結果に基づき、効果モデルの改定案を整理した後、再度、実践家参画型ワークショップを実施し、第2次効果モデル（提示版）を構築した【ステップ10に該当】。さらに、実施組織の効果モデル実施段階を測定するためのフローチャート（実施組織の効果モデル導入

段階がアセスメントできる)、戦略プラン(実施組織の強みや弱み、目標を明らかにできる)、実施プラン(実施組織の効果モデル導入の具体的な計画を策定できる)等の作成も行った(効果のあがる就労移行支援プログラムのあり方研究会 2017)。

## ②評価キャパシティ形成( ECB)の取組み

このようなプログラムモデル(効果モデル)構築のプロセスは同時に、実践家の評価キャパシティ形成( ECB)も兼ねていた。

まず、取組み当初に実施したGP( Good Practices)事例調査の対象になった実践家とは、その後のフォーカス・グループインタビュー、第1次効果モデル(試行版)構築の作業を協働した(視点①)。このように、第1次効果モデル(試行版)構築までの作業を協働することによって、実践家のプログラムモデル(効果モデル)に関する主体性は増していき、この取組みへのより積極的な関与を引き出すことが可能になった(視点②)。また、プログラムモデル(効果モデル)構築の作業を実践家と協働できたことによって、実践家からの多くのインプットを得ることができ、実施組織に受け入れられやすい、現実的なプログラムモデル(効果モデル)の構築を行うことができた(視点③)。

その後、ここまでの取組みを協働してきた実践家には、2009年2月～2009年12月にかけて実施した全国試行評価調査①(単一グループで行う多施設共同調査)で「実践家評価担当者」の役割、すなわち、それぞれの実施組織においてプログラムモデル(効果モデル)実施の中心的な役割、実施組織における評価(プロセス評価・アウトカム評価)の中心的な役割を依頼した。なお、この調査を開始する前には、実践家評価担当者を対象にしたプログラムモデル(効果モデル)に関する研修会を実施した(視点①)。

さらに、1度目の全国試行評価調査①にて「実践家評価担当者」の役割を担った実践家の多くは、2013年11月から2014年11月にかけて実施した2度目の調査においては、「実践家評価ファシリテーター」の役割を担当し、他の実施組織に所属する実践家評価担当者へのコンサルテーショ

ン等を行った(視点②)。なお、このときは「実践家評価担当者・評価ファシリテーターマニュアル(第2版)」(効果のあがる就労移行支援プログラムのあり方研究会 2015)を作成し、「実践家評価担当者」向けの研修会に加えて、「実践家評価ファシリテーター」を対象にした研修会も実施した(視点①)。そして全国試行評価調査①のなかで行った実施組織の訪問調査(プロセス評価)では研究者評価ファシリテーターが実践家評価ファシリテーターのコンサルテーションを行い、実践家参画型エンパワメント評価(CD-TEP法)の観点からサポートした(視点①)。

このように、全国試行評価調査①の取組みを協働することによって、当初から協働してきた実践家らは、実践家評価担当者から実践家評価ファシリテーターへと役割を変え、よりこの取組みへの主体性及び関与を高め、さらには実践家参画型エンパワメント評価(CD-TEP法)への理解も深めていった(視点②)。また、実践家評価担当者にとっても同じ実践家という立場を有する実践家評価ファシリテーターからのコンサルテーションを受けられたことで、より効果的にプログラムモデル(効果モデル)への理解を深めることができ、さらにはプログラムモデル(効果モデル)の実施度(フィデリティ尺度得点)のより良い向上にも貢献した(視点②)。こうした実践家の積極的・主体的関与は、全国試行評価調査①をより充実したものにし、プログラムモデル(効果モデル)の改訂に資する多くの有益な情報を得ることに繋がった(視点③)。

最後に、このプログラムを実施・普及するためのフローチャートや戦略・実施プランを備えたことは、実践家評価ファシリテーターのコンサルテーションをより容易に、そして有効にした(視点①・②)。なぜなら、これらは実践家評価担当者やその他のプログラムスタッフがプログラムモデル(効果モデル)導入に向けて、実施組織の準備性を高めることのみならず、実践家評価ファシリテーターがコンサルテーションを行う際にも有用に機能するからである。

### (3) 生活困窮者自立相談支援事業プログラム

#### ①プログラムモデル（効果モデル）構築のプロセス

生活困窮者自立相談支援制度にあるいくつかの事業は任意事業の扱いになっており、結果として、自治体（実施組織）ごとに取組まれ方も様々、その結果、成果の出方も様々という問題を抱えていた（新藤ら 2018）。そこで筆者らの研究チームはCD-TEP法を用いてこの事業の効果モデル構築に取組むこととした。

効果モデル構築の取組みは2017年9月に開始され、まず、全国の先進的な実践事例に取組み、成果を上げている実践家数名と一緒にロジックモデルづくりの参加型ワークショップを実施した。その後、さらに複数のGP（Good Practices）事例を対象に訪問調査を実施し、この事業に取組むうえでの具体的な工夫について、聞き取りを行った【ステップ1～3に該当】。その後、ここまでのデータをもとにロジックモデルと効果的援助要素リスト、アウトカムを測定するための評価ツール、フィデリティ尺度を作成し、第1次効果モデル（試行版）を構築した【ステップ4～6に該当】。2017年12月には、全国1,317か所の自立相談支援機関を対象に、第1次効果モデル（試行版）に基づく広域的事業所調査（アンケート調査）を実施し、フィデリティ尺度得点（効果的援助要素の実施）がアウトカム指標と正の相関関係にあることを確認した【ステップ7・9に該当】（新藤ら 2018）。なお、ここまでの研究成果は厚生労働省の担当者や多くの実践家等が参加したセミナーを開催して報告し、厚生労働省の担当者やこの事業に取組む実践家らとの情報共有を図った（新藤ら 2021）。その後、2018年10月から11月にかけて、再度、GP（Good Practices）事例調査を実施し、ここで得たデータをもとに効果的援助要素リストの改訂を行った【ステップ3～6に該当】。そして、2018年12月から2019年3月にかけて、全国7機関を対象にした全国試行評価調査①（単一グループで行う多施設共同調査）を実施し、フィデリティ尺度得点（効果的援助要素の実施度）の上昇にともなって、アウトカム（利用者の就労意欲や様々なスキル、就労者数等）も向上していることを確認した（ユ

ニバーサル志縁センター 2019）【ステップ8・9に該当】。ここまでの研究成果も、厚生労働省の担当者や多くの実践家等が参加したセミナーを開催して報告し、厚生労働省の担当者やこの事業に取組む実践家らとの情報共有を図った。また、こうして構築した効果モデルは、生活困窮者自立支援全国ネットワークが開催する九州・沖縄ブロック研修会でも取り上げられ、多くの実践家らにこのプログラムモデル（効果モデル）の内容を伝えるための機会を得た（新藤ら 2021）。

#### ②評価キャパシティ形成（ECB）の取組み

この事例においても、プログラムモデル（効果モデル）構築のプロセスは、同時に、評価キャパシティ形成（ECB）を兼ねた。

まず、取組み当初に実施したロジックモデルづくりの参加型ワークショップは、これに参加した実践家のこの取組みへの関心・主体性を高めた。現在、この実践家のうちの一人は様々な場所で自主的なワークショップを開催しているが、これに関して「このときのロジックモデルづくりの経験が大きかった」と話している（視点①・②）。また、作成されたロジックモデル自体も、ワークショップを通して実践家と作業を協働し、実践家の日々の実践から多くのインプットを得たことで、多くの実施組織に受け入れられやすい、現実的なモデルの構築を行うことができた（視点③）。

その後、全国1,317か所の自立相談支援機関を対象にした広域的事業所調査（アンケート調査）に進むが、この調査によって得られたデータの分析結果を、実践家らと共有しながら、取組みを進めたため（視点①）、実践家らのCD-TEP法（データ収集方法やデータ分析方法）に関する理解もより深まった（視点②）。

また、定期的に厚生労働省の担当者や情報共有を行いながら取組みを進捗させ、セミナー開催を通して成果の共有も図りながら取組みを進めたため（視点①）、厚生労働省の担当者からも肯定的なフィードバックを得ることができ、制度的にも、良い方向に制度改正がなされ、現状の課題解決が少しずつ進むかもしれないという期待感を得ることができた（視点③）。こうした

厚生労働省との関係とその成果は、実践家らのこの取組みに対するモチベーションを高く維持することにも繋がった（視点②）。

さらに、当初のロジックモデルづくりワークショップに参加した実践家の一人は、2018年12月から2019年3月にかけて実施した全国試行評価調査①において、実践家評価ファシリテーターの役割を担い、調査対象となった7機関の実践家評価担当者及びその他のプログラムスタッフに対して、効果モデルに基づく実践現場のコンサルテーションを行った（視点②）。実際のコンサルテーションでは、アクション・プランシートという実施組織のアセスメント・計画策定ツール（戦略プラン・実施プランと同じ機能を持つツール）を使用して、対象実施組織の効果モデル導入に関する強みや弱みを把握し、これらを踏まえた効果モデル導入のための現実的な目標設定を行ったうえで、実践家評価ファシリテーターからのコンサルテーションが行われた（視点①）。こうした実践家評価ファシリテーターからのコンサルテーションは、実践家評価担当者やその他のプログラムスタッフのプログラムモデル（効果モデル）への理解を深め、モチベーションを高めた。さらに、このことは実施組織におけるフィデリティ尺度得点（効果的援助要素の実施度）を高めることにも貢献した（ユニバーサル志縁センター 2019；新藤ら 2021）。ここまでの研究プロセス、研究成果も厚生労働省の担当者と共有しながら進捗させたため、前述の効果を継続して得ることができている（視点①・②）。

現在、広域的試行評価調査①にて実践家評価ファシリテーターの役割を担った実践家は、この取組みへの主体性・関与度をさらに高め、「EBP効果モデル」技術支援センターの一員として継続した取組みへの協力とともに、様々な実施組織を対象にした自主的なワークショップ、本プログラムモデル（効果モデル）の実施・普及活動を行ってくれている（視点②）。

## 4. 総合考察

以上2事例の分析結果を踏まえて、本稿の目的であった評価キャパシティ形成（ECB）に果たす「EBP効果モデル」技術支援センター（EBP-TAC）の役割について、まずは、大島ら（2019、p.255）が作成した「EBP効果モデル」技術支援センター（EP-TAC）のロジックモデル（図2）は本稿に示した2事例の実践に当てはまるかを検証する。次に、これを踏まえたうえで評価キャパシティ形成（ECB）に関して、「EBP効果モデル」技術支援センター（EBP-TAC）が果たすべき役割とは何かを考察する。そして最後に、「EBP効果モデル」技術支援センター（EBP-TAC）への期待と今後の課題を述べる。

### (1) 「効果モデル」技術支援センター（EBP-TAC）のロジックモデルと2事例の対照

前述の通り、大島ら（2019、p.255）が作成した「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデル（図2）では、評価キャパシティ形成（ECB）に関連したアウトカムを「実施アウトカム」、「サービス・アウトカム」、「利用者アウトカム」の3段階から捉えている。

本稿で示した「効果的障害者就労移行支援プログラム」及び「生活困窮者自立相談支援事業プログラム」の事例から、とりわけ「評価キャパシティ形成（ECB）」の取組みを整理し、大島ら（2019、p.255）が作成した「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデルに当てはめたものが図3に示した「EBP-TACのロジックモデルと2事例の対照」である。

まず、2事例では共通して研究者評価ファシリテーターと実践家が協働して効果モデル構築の作業に取り組んでいたことが確認された。このことは、協働した実践家評価担当者の取組みに対する積極的な関与を引き出すことにつながり、また効果モデルやその評価方法に関する実践家評価担当者の理解を深めることにも貢献した。次に、2事例は共通して戦略プラン・実施プラン（アクション・プランシート）を活用して効果モデルの実施状況をアセスメントし、この取組みをどう進めていけば良いのかに関する計画づく

図3 EBP-TACのロジックモデルと2事例の対照

EBP-TAC 支援プロセス	実施アウトカム	サービス・アウトカム	利用者アウトカム
<p>①協働による効果モデル開発の取組み《評価支援》</p> <ul style="list-style-type: none"> <li>・効果モデル構築の作業を協働</li> <li>・データ分析の結果を実践家らと共有しながら取組みを進める</li> </ul> <p>②アセスメントと評価キャパシティ形成のための計画策定《効果モデル実施支援・評価支援》</p> <ul style="list-style-type: none"> <li>・戦略プラン・実施プランによるアセスメントと効果モデル実施、評価に関する計画の策定</li> </ul> <p>③具体的な支援メニューの提供《効果モデル実施支援・評価支援》</p> <ul style="list-style-type: none"> <li>・実践家評価担当者・実践家評価ファシリテーターマニュアル作成</li> <li>・実践家評価担当者及び実践家評価ファシリテーターを対象にした効果モデルに関する研修会の実施</li> <li>・研究者評価ファシリテーターによる実践家評価ファシリテーターへのコンサルテーション</li> </ul> <p>④ネットワーキング《組織づくり》</p> <ul style="list-style-type: none"> <li>・厚生労働省の担当者と情報を共有しながらのセミナー開催</li> </ul>	<ul style="list-style-type: none"> <li>・実践家評価担当者の積極的な関与が引き出される</li> <li>・実践家のデータ収集方法やデータ分析方法に関する理解が深まる</li> <li>・実践家評価担当者及びプログラムスタッフの効果モデルに関する理解が深まる</li> <li>・多くの実践家が実践家評価ファシリテーターとして活躍</li> <li>・厚生労働省の担当者から良いフィードバックが得られる</li> </ul>	<ul style="list-style-type: none"> <li>・効果モデルの実施度（フィデリティ尺度得点）が向上</li> <li>・多くの実践家評価担当者が効果モデルを適切に理解</li> <li>・実践家から多くのインプットを得たことによる効果モデルの改善</li> <li>・実践家らの取組みに対するモチベーションが向上</li> <li>・実践家評価ファシリテーターは積極的に効果モデルの普及活動に取組む</li> </ul>	<p>利用者アウトカムの向上 (2つの事例に共通したアウトカムはプログラム受益者が就労すること)</p>

(出所) 大島ら (2019, p.255) をもとに筆者作成

りを行っていた。こうした取組みは実践家評価ファシリテーターの効果モデル実施支援や評価支援を行いやすくし、また、実践家評価担当者及びその他のプログラムスタッフにとっても、実施組織の現状に見合った支援が受けられることから、より一層、効果モデルに関する理解を深めることに貢献した。また、実践家評価担当者及び実践家評価ファシリテーターに対して、実践家参画型エンパワメント評価（CD-TEP法）への取り組み方に関するマニュアルを作成し、これをもとにした研修を実施したこと、さらに

実践家評価ファシリテーターに対しては研究者評価ファシリテーターからのコンサルテーションがあったことは、実践家の関心と理解を引き出すことにつながり、実践家評価ファシリテーターとして取組みに参加する実践家を増やしていくことに貢献した。

最後に、生活困窮者自立相談支援事業プログラムの事例では、実践家参画型エンパワメント評価の取組みを厚生労働省の担当者にも共有しながら進め、セミナーの開催等を通して意見交換、研究で得た知見の発信に取組んだことで、厚生

労働省担当者の関心を高め、取組みに対する良いフィードバックを得ることにつながった。

こうした取組みは大島ら（2019、p.255）が作成した「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデル（図2）に示した活動におおよそ該当するものであり、またこのような活動から得た成果は、同じく図2の「実施アウトカム」に示された「組織の実施」、「スタッフの実施態度」、「実施システムの持続可能性」の向上に該当するものであったと考えられる。また、こうした「実施アウトカム」の達成が、その先の「サービス・アウトカム」、「利用者アウトカム」の向上につながることも確認されており、大島ら（2019、p.255）が作成した「EBP効果モデル」技術支援センターが行う評価キャパシティ形成（ECB）のロジックモデル（図2）が実践事例との対照からみても一定程度妥当であることが確認されたといえる。

## (2) 「EBP効果モデル」技術支援センター（EBP-TAC）がECBに果たすべき役割

このように、本稿における2つの事例分析を通して、実践事例という側面からも「EBP効果モデル」技術支援センター（EBP-TAC）が行うべき評価キャパシティ形成（ECB）支援が明らかになった。これを一定の手順で示すとすれば、①協働による効果モデル開発の取組み、②アセスメントと評価キャパシティ形成（ECB）のための計画策定、③具体的な支援メニュー（トレーニング・コンサルテーション・見学等の機会）の提供、④ネットワークに整理できる。なお、図2に示した「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデルの中で、とりわけ支援プロセス（図中、左側の列）の部分に当てはめると、このうち②と③は「効果モデルの実施支援」及び「評価支援」の両方に、①は「評価支援」に、④は「組織づくり」に該当すると考えられる（図3）。

実践家、あるいは実施組織の評価キャパシティ形成（ECB）において、まずは、①協働による効果モデル開発の取組み、そのための評価活動への実践家の積極的な参画が欠かせない。本稿で取り上げた2事例においては、いずれも最初の

段階で効果モデルのインパクト理論・プロセス理論、あるいはロジックモデルづくりへの実践家の十分な関与を得ることができ、このことが実践家参画型エンパワメント評価（CD-TEP法）への継続した関与を引き出したことが確認された。このことは、利害関係者が評価活動に参加することで、当該プログラム及び評価活動への主体性を高め、評価に関する知識や技術も高めていくというエンパワメント評価の理論（Wandersman et al.（=2014））からも裏付けを得ることができる。もし、「EBP効果モデル」技術支援センター（EBP-TAC）によるこのような取組みがなければ、日々自身の業務で多忙を極める実践家らは、実践家参画型エンパワメント評価（CD-TEP法）の取組みに対するモチベーションを維持することが難しく、実践家からの継続した取組みへの参加はいずれ得られなくなってしまうことが想定される。

次に行うべきは、②アセスメントと評価キャパシティ形成（ECB）のための計画策定である。効果的障害者就労移行支援プログラムの事例では効果モデル実施・普及のためのフローチャート・戦略プラン・実施プランを備え、生活困窮者自立相談支援事業プログラムの事例ではアクション・プランシートを用いた。生活困窮者自立相談支援事業プログラムの事例で、実践家評価ファシリテーターが対応したように、「EBP効果モデル」技術支援センター（EBP-TAC）は、実施組織におけるプログラムモデル（効果モデル）導入状況、その準備状況をアセスメントするとともに、実践家評価担当者やその他のプログラムスタッフが実践家参画型エンパワメント評価（CD-TEP法）の取組みをどのように捉えているのか、何か困難が生じていないか、どのようなサポートを必要としているか、等をアセスメントする必要がある。「EBP効果モデル」技術支援センター（EBP-TAC）は、こうしたアセスメントの結果に基づき、実施組織（実践家評価担当者やその他のプログラムスタッフ）に役立つ妥当な計画を策定したうえで、計画的に評価キャパシティ形成（ECB）の取組みを進めていくことが有効である。もし、「EBP効果モデル」技術支援センター（EBP-TAC）によるこのような

取組みがなければ、効果モデルを実施しようとする実施組織にとって的外れなアドバイスや支援をしてしまうことも想定され、このことは実践家の効果モデル及び評価に対する理解の促進を妨げ、やはりいずれ実践家からの継続した取組みは得られなくなってしまうだろうと考えられる。

次に、評価キャパシティ形成（ECB）のための具体的な計画を策定するために、あるいは計画策定後、実際に評価キャパシティ形成（ECB）の取組みを進めていくために、「EBP効果モデル」技術支援センター（EBP-TAC）は、③評価キャパシティ形成（ECB）のための具体的な支援メニュー（トレーニング・コンサルテーション・見学等の機会）を備えている必要がある。本稿で取り上げた2つの事例はいずれも、広域的試行評価調査①（単一グループで行う多施設共同調査）を開始するタイミングでプログラムモデル（効果モデル）及び評価活動に関する研修会を開催している。また、広域的試行評価調査①においては、実践家評価ファシリテーターからのコンサルテーションが行われ、これが実践家評価担当者やその他のプログラムスタッフの評価キャパシティ形成（ECB）に有効であったことも確認されている。なお、実践家評価ファシリテーターに対しては、研究者評価ファシリテーターからのバックアップ支援が提供された。これは実践家評価ファシリテーターを対象にした評価活動に関するトレーニングの一環であったといえる。その他、実践家同士による実施組織のピア訪問（見学会）等も有効であると考えられる。このように、「EBP効果モデル」技術支援センター（EBP-TAC）は評価キャパシティ形成（ECB）に関するいくつかの支援メニューを備え、実践家及び実施組織のニーズに応じて、適切なものを提供する必要がある。なお、支援メニューの選択は、評価キャパシティ形成（ECB）に関する実践家あるいは、実施組織のアセスメント及び計画策定にて行われることは前述の通りである。例え、戦略プラン・実施プランに基づいて、実施組織の効果モデル実施あるいは評価に関する計画が立てられたとしても、「EBP効果モデル」技術支援センター（EBP-TAC）がそれを実施す

るための具体的な支援メニューを持っていないければ、実施組織の助けになることはできない。そのため、まずは「EBP効果モデル」技術支援センター（EBP-TAC）が存在していることが重要だが、そのうえで、これが十分な支援メニューを備えていることも、また重要なことであるといえる。

最後に、実践家参画型エンパワメント評価の取組みを進めていくための④ネットワーキングを挙げる。本稿で取り上げた2事例は、いずれも実践家参画型エンパワメント評価（CD-TEP法）の取組みを進め、プログラムモデル（効果モデル）を構築していくなかで、この取組みに参加する実践家等関係者のネットワークを拡大してきた。こうして拡大していくなかで、実践家評価ファシリテーター、実践家評価担当者及びその他のプログラムスタッフと異なる役割を担う実践家が生まれ、それぞれが取組むべき活動を担うなかで、それぞれの評価キャパシティ形成（ECB）も進捗した。このことに加えて、生活困窮者自立相談支援事業プログラムの事例では、厚生労働省担当者との繋がりも継続しながら実践家参画型エンパワメント評価（CD-TEP法）の取組みを進め、このことがさらに実践家のこの取組みへのモチベーションを高めることに貢献した。このようなネットワーキングの取組みは評価キャパシティ形成（ECB）の観点からも、実践家参画型エンパワメント評価自体の観点からも重要な取組みであり、「EBP効果モデル」技術支援センター（EBP-TAC）が担うべき役割の一つであるといえる。本稿冒頭で述べた通り、実践家参画型エンパワメント評価は普遍的・効果的なプログラムモデル（効果モデル）の開発を目指している。そのため、より多くの実施組織の参画、厚生労働省のような制度をつくる立場の人の参画が欠かせない。これは研究者や実践家が個人的に行う活動では大きな成果を得ることが難しく、やはり「EBP効果モデル」技術支援センター（EBP-TAC）のような中間支援組織の存在が重要になると考えられる。

### (3) 「EBP効果モデル」技術支援センター（EBP-TAC）への期待と今後の課題

以上を踏まえて、最後に、「EBP効果モデル」技術支援センター（EBP-TAC）への期待と今後の課題について述べる。本稿冒頭で述べたとおり、近年、形成的評価やプログラム開発に対する関心が高まっている。例えば、2010年にイギリスではじまった行政の成果連動型支払い契約と民間資金の活用を組み合わせた官民連携手法の一つであるソーシャル・インパクト・ボンドの試行プロジェクトが2015年に日本でも開始され、その後、いくつかの実証事例や導入事例が生まれている。また、おおよそ同時期から日本で社会的インパクト評価が普及しはじめ、現在では、社会的インパクト・マネジメント・イニシアチブのもとにその積極的普及が図られ、2019年には社会的インパクト評価を仕組みとして位置付けた休眠預金活用事業が開始された（一般財団法人日本民間公益活動連携機構 2019）。

ソーシャル・インパクト・ボンドにしても、社会的インパクト評価にしても、これらに取組むうえで求められるのは、新しく効果的なプログラムを開発し、より有効なものへと改善するための理論と方法論である。本稿で提示した実践家参画型エンパワメント評価はそのための有効な方法論の一つであり、今後、この方法論のより一層の普及と、様々なプログラム領域での普遍的・効果的なプログラムモデル（効果モデル）の構築が求められる。同時に、「EBP効果モデル」技術支援センター（EBP-TAC）の実装・設置を進め、評価キャパシティ形成（ECB）を通じたより多くの評価人材の育成も望まれる。

そのためには、「EBP効果モデル」技術支援センター（EBP-TAC）に関する研究を進め、より一層、果たすべき役割、効果的な運営方法を明確にしていく必要がある。特に、現状、日本における「EBP効果モデル」技術支援センター（EBP-TAC）の取組みは、研究ベースで（国等の研究資金によって）実施されることがほとんどである。大島ら（2019、p.358）は、公的な資金投入がないなかでも「EBP効果モデル」技術支援センター（EBP-TAC）の運営は可能と述べているが、どのような運営の仕方であれば、それが

無理なく実現できるのか、等については今後、明らかにしていかなければならない大きな課題の一つであると考えられる。

## 5. 結論

本稿では、これまでに筆者が取組んできた効果的障害者就労移行支援プログラムの事例及び生活困窮者自立相談支援事業プログラムの事例をもとに、実践家参画型エンパワメント評価の基盤を支える「EBP効果モデル」技術支援センター（EBP-TAC）の意義と役割について、評価キャパシティ形成（ECB）の観点から考察した。

その結果、まずは、大島ら（2019、p.255）が作成した「EBP効果モデル」技術支援センター（EBP-TAC）のロジックモデル（図2）は実践的側面からみても一定程度妥当であることが確認された。また、このロジックモデルにもあるように、「EBP効果モデル」技術支援センター（EBP-TAC）が取組む評価キャパシティ形成（ECB）のアウトカムには「実施アウトカム」、「サービス・アウトカム」、「利用者アウトカム」の3段階があることを確認した。次に、こうした評価キャパシティ形成（ECB）に関連した成果（アウトカム）を得るための働きかけの手順として、①協働による効果モデル開発の取組み、②アセスメントと評価キャパシティ形成（ECB）のための計画策定、③具体的な支援メニュー（トレーニング・コンサルテーション・見学等の機会）提供、④ネットワーキングがあることを確認した。

今後、実践家参画型エンパワメント評価の普及と、「EBP効果モデル」技術支援センター（EBP-TAC）の実装・設置が望まれるが、そのためにも、より一層の「EBP効果モデル」技術支援センター（EBP-TAC）の役割の明確化、有効な運営方法の検討等が必要になるだろう。

## 謝辞

効果のあがる就労移行支援プログラムの取組みを共に進めていただいた研究班の皆様、生活困窮者自立相談支援事業プログラムの取組みを

共に進めていただいたユニバーサル志縁センターの池本修悟様、A'ワーク創造館の西岡正次様、それぞれの研究にご協力いただきました関係機関の皆様にご心よりお礼申し上げます。

## 付記

本研究は、文部科学研究費基盤研究（A）（課題番号：15H01974・23243068）（研究代表者：大嶋巖）及び平成29年度・30年度厚生労働省社会福祉推進事業（事業実施主体：ユニバーサル志縁センター）の一部として行われ、本研究の一部は、日本評価学会第20回全国大会（2019年12月）、日本ソーシャルワーク学会2021年度第38回大会（2021年7月）で発表した。

## 参考文献

一般財団法人民間公益活動連携機構（2019）『資金分配団体・実行団体に向けての評価指針』、一般財団法人日本民間公益活動連携機構

大嶋巖・源由理子・山野則子・賛川信幸・新藤健太・平岡公一（2019）『実践家参画型エンパワメント評価の理論と方法—CD-TEP法：協働によるEBP効果モデルの構築』、日本評論社

効果のあがる就労移行支援プログラムのあり方研究会（2015）『実践家評価担当者・評価ファシリテーターマニュアル（第2版）』、文部科学省科学研究費補助金基盤研究（A）実践家参画型効果的プログラムモデル形成評価研究班

効果のあがる就労移行支援プログラムのあり方研究会（2017）『実施・普及マニュアル（第3次全国試行評価調査配布版）』、文部科学省科学研究費補助金基盤研究（A）実践家参画型エンパワメント評価を活用した有効なEBP技術支援センターモデル研究班

新藤健太・大嶋巖・浦野由佳・植村英晴・方真雅・村里優・全形文（2017）「障害者就労移行支援プログラムにおける効果モデルの実践への適用可能性と効果的援助要素の検討～全国22事業所における1年間の試行的介入研究の結果から」、『社会福祉学』、58（1）：57-70

新藤健太・源由理子・鴨崎貴泰・大澤香織（2018）「自立相談支援事業（生活困窮者自立支援制度）にお

る事業評価ガイドラインの開発～セオリー評価・プロセス評価に注目して」、『日本評価学会春季第15回全国大会抄録集』

新藤健太・池本修悟・池田徹・源由理子（2021）「生活困窮者自立相談支援事業に導入された実践家参画型エンパワメント評価の成果と、その成果生成を支える中間支援組織の機能と役割」、『日本ソーシャルワーク学会2021年度第38回大会資料』

ユニバーサル志縁センター（2019）『自立相談支援事業評価実践ガイド普及展開方法検討事業報告書』、平成30年度厚生労働省社会福祉推進事業報告書

Institute of Education Sciences(IES): National Center for Education Evaluation and Regional Assistance(2011). *National Evaluation of the comprehensive technical assistance centers: Final report*. IES.

Labin, SN., Duffy, JL., Meyers, DC, Wandersman, A and Lesesne, CA.(2012). A Research Synthesis of the Evaluation Capacity Building Literature. *American Journal of Evaluation*, 33(3), 307-338.

Preskill, H. and Boyle, S.(2008). Insights into Evaluation Capacity Building: Motivations, Strategies, Outcomes and Lesson Learned. *The Canadian Journal of Program Evaluation*, 23(3), 147-174.

Proctor, EK., Powell, BJ. and McMillen, JC.(2013). Implementation strategy: Recommendations for specifying and reporting. *Implementation Science*, 8, 139-150.

Wandersman, A. and Fetterman, DM.(2005). *Empowerment Evaluation Principles in Practice*, Guilford Press. (=2014, 笹尾敏明監訳・玉井航太・大内潤子訳『エンパワメント評価の原則と実践—教育、福祉、医療、企業、コミュニティ介入プログラムの改善と活性化に向けて』、風間書房)

(2022.6.10 受理)

**The Significance and Role of the EBP (Evidence-Based Practices)  
Technical Assistance Center, the Foundation for Practitioner-Based  
Empowerment Evaluation  
—Focusing on the Evaluation Capacity Building—**

Kenta Shindo

Japan Collage of Social Work  
shindo@jcsww.ac.jp

Iwao Oshima

Tohoku Fukushi University  
i-oshima@nifty.com

**Abstract**

Practitioner-Based Empowerment Evaluation (CD-TEP) approach is a useful method of program development and evaluation, aiming to obtain, implement, and disseminate effective program models. As part of promoting this approach, it is essential to build the evaluation capacity of the practitioners and organizations engaged in these programs. In this context, we examined the significance and role of the EBP (Evidence-Based Practices) Technical Assistance Center, which supports the building of evaluation capacity.

The results indicate that evaluation capacity building leads to a change within practitioners and organizations, and even to an improvement of service outcomes or client outcomes. Accordingly, it appears that an effective EBP (Evidence-Based Practices) Technical Assistance Center performs (1) development of effect model through collaboration, (2) planning for assessment and evaluation capacity building, (3) provision of opportunities of training, consultation, observation, etc., and (4) networking.

**Keywords**

EBP Technical Assistance Center, Evaluation Capacity Building, Practitioner-Based Empowerment Evaluation

## 【研究論文】

# 自治体会計情報のオープンデータ化の促進 —国際基準 COFOG による費目分類に着目して—

生方 裕一

早稲田大学

y1\_ubukata@aoni.waseda.jp

佐藤 亨

日本生産性本部

toru.sato@jpc-net.jp

川島 宏一

筑波大学

hkawashima@sk.tsukuba.ac.jp

## 要 約

本稿は、日本の地方自治体が会計情報や財務書類を用いて説明責任を果たす際、会計情報の整備に注力するあまり、情報利用者、特に住民に対する視点が不足している点を指摘するものである。欧米では、住民視点で地方自治体の予算を、COFOGと呼ばれる分類基準を用いて分かりやすく可視化する取り組みが行われている。本稿では、このCOFOGを用いることで住民目線に立つ会計情報が作成され、各自治体の会計情報の説明責任、および透明性の向上に寄与する可能性があることを考察する。その上で、今後の公会計改革への示唆を得ることを目的とする。今後、日本の地方自治体において、説明責任を果たすための公会計情報の整備が求められることを踏まえるとCOFOGに基づいて、住民視点に立つ細かな会計情報を用意することが必要となろう。このような会計情報が用意されることで、住民は、自治体運営の全体像を容易に理解できるようになる。さらには、住民の行政への参画を促進させることも期待される。

## キーワード

地方自治体、会計情報、オープンデータ、COFOG

## 1. はじめに

1990年代以降、財政状況の改善等を図るべく、世界各国の公的組織においてNew Public Management (NPM) に立脚した行政活動が行われてきた。昨今では、公的部門における説明責任の履行と会計情報等の各種情報を公開し、組織の透明性を高めるために、New Public Governance (NPG) に立脚した行政活動が行われる傾向にある (Bryson et al. 2014; Fimreite & Læg Reid 2009)。

さらに2009年にオバマ米大統領 (当時) が、Open Government (開かれた行政組織)、

Transparency (行政組織の透明性)、Collaboration (官民の連携) からなる3つの原則を提示して以降、透明性を高め「開かれた行政組織」を目指す取り組みが世界中で行われている。具体的には、住民が情報機器等を活用した行政組織の文書等の入手や、各種手続きを行う権利等を保障することで、行政組織の透明性を高める取り組みが見られる。この傾向は、住民目線に立つ財務報告 (Popular Financial Reporting) の普及にも表れている (Steccolini 2004)。この財務報告は、従来の財務報告の欠点を補いつつ (Steccolini 2004)、会計情報自体の透明性を高め、住民に対し、行政への参画を促進させる (Biondi & Bracci

2018; Manes-Rossi 2019; Steccolini 2004)。これは、住民目線の財務報告によって、行政組織と住民間の情報共有が円滑になることで、情報自体の透明性が高まり、行政活動への住民参画が促されるからである (Célérier & Cuenca 2015; Manes-Rossi et al. 2019)。

日本の地方自治体における財務報告は、内閣による国会報告<sup>1</sup>や地方財政白書の発行等という形で行われている。さらに昨今では、公会計改革により、発生主義会計情報およびそれに基づく財務書類の整備<sup>2</sup>が進められている。これにより、地方自治体内外の利害関係者に対する財政状況の報告がなされ、収支バランスの把握と財政健全化に向けた財政の統制が期待されている。

このような統制への期待は、地方自治体が公開した公会計情報の利活用が前提である。しかし、大多数の地方自治体では、固定資産台帳の整備と統一的な基準に基づく財務諸表の期限内の準備に留まる (山浦 2016) というのが現状である。そのため、会計情報に基づいた説明責任の履行については一定の成果が見られるものの、それ以外の点での活用が課題である (東 2020)。具体的には、適切な資産管理、予算編成への活用、政策・施策・事務事業等の評価との連携等、政策会計情報の整備があまり進んでいない (総務省 2015)。

上述の課題に関して、本稿では説明責任を果たす上で、会計情報の整備に注力するあまり、情報利用者<sup>3</sup>、特に住民に対する視点が不足している点を指摘したい。すなわち「予算制度とその具体的現れとしての予算書が住民の意思決定を明瞭に示し、意思決定に十分参加しえる情報を提供しているかどうか」(山本 1991, p.4) という視点である。この視点は、当該会計情報を、住民が理解し、活用できるという点で有益であるかどうかを問うものである。

筆者らは、住民にとって有益な会計情報を「自分が納めた税金がどのような政策等に使われているのかをわかりやすく示したもの」と想定している。わかりやすい会計情報の提示については、英国や欧州諸国の事例が参考になる。これらの国々では、特に地方自治体において、当該地域の住民の手により、行政組織による税金の

使われ方をより深く理解するために、当該地方自治体の予算を住民視点で分かりやすく可視化する取り組みが行われている。この取り組みの特徴的な点として、地方自治体の予算の可視化にあたり、国連が基準を定義したCOFOG (Classification of the Functions of Government, 機能別分類) で予算費目を分類している点である。COFOGによる分類で、わかりやすく可視化されるとともに、国内外の同規模地方自治体との比較可能性も高めている。

さて、日本の地方自治体の場合、現行の予算からでは、COFOGに基づく分類は難しい。なぜならば、日本の地方自治体は、地方自治法に基づく「款・項・目・節」という費目区分で予算が区分されているからである。そこで本稿では、各自治体で整備されている会計情報、特に会計情報の基礎となる予算の費目区分が、当該自治体財政状況について住民の理解を妨げている可能性を指摘したい。そして、会計情報の理解を促す費目区分として、COFOGによる分類を提案する。これは、COFOGに基づいて分類された会計情報が住民目線に立つ会計情報となり、各自治体の会計情報の説明責任、および透明性の向上に寄与する可能性があるためである。そして、今後の公会計改革への示唆を得ることを目的とする。

## 2. 地方自治体における予算の費目分類

### (1) 予算の費目分類

地方自治体に限らず、行政組織における予算の記録・表示・報告方法は、予算の費目ごとの分類に応じて決定される (Jacobs, Héris & Bouley 2009)。これは、政策立案者、政策実施担当者、行政組織、議会、及び住民等に正確な情報を提供するためには、予算の費目分類が必要だからである (Jacobs, Héris & Bouley 2009)。

この分類により、当該行政組織の予算管理の、健全な予算管理に必要な3つの主要原則<sup>4</sup> (包括性の原則・単一性の原則・内部一貫性の原則) への準拠が確認される。そして、予算が行政組織の総支出および部門別支出を抑制し、資源配分

の効率化を促進する上で効果的であることが保証される (Jacobs, Héris & Bouley 2009)。さらに、予算の費目分類により、支出と収入が正しく分類される。その結果、政策決定と説明責任の両方に規範的な枠組みが提供され、政策立案とパフォーマンス分析の促進、セクター間での効率的な資源配分、立法府が承認した予算資源の遵守の確保、予算の透明性の保証、日々の業務の改善がなされる (Jacobs, Héris & Bouley 2009)。以上のことから、健全な予算管理には、健全な費目分類が必要である (Jacobs, Héris & Bouley 2009)。

## (2) 「款・項・目・節」の導入経緯

日本の地方自治体の業務は、住民の福祉の維持向上を目的として、住民が安全・快適に生活できるようにすることである (井上 2005)。この業務を遂行するために、各地方自治体は、会計情報等、自治体行政に関する数値情報を記録・計算・報告する (井上 2005)。

そもそも会計とは「組織体の活動を貨幣額で記録、計算、報告すること」(井上 2005, p.94)であり、その手法や形態が「組織体の目的に従う」(井上 2005, p.94)ものである。加えて、その会計が主に誰に対しての報告であるのかという視点が重要となる (井上 2005)。地方自治体の目的が「住民の福祉の維持向上」にあることを踏まえると、地方自治体の会計は「住民に対して自治体の財政状況を正しく、わかりやすく報告する仕組みと形式を提供するもの」(井上 2005, p.94)といえよう。したがって、地方自治体会計制度の意義と役割は「自治体の行政活動を反映した財政状況について、住民に対し、その知る権利に十分応えられるような会計報告を作成し、報告すること」(井上 2005, p.95)にある。以上のことから「自治体が住民の福祉の増進のための行政活動を行うための予算とその執行実績である決算は、その目的の達成度合いがわかりやすく、正しく示されることが必要」(井上 2005, p.94)である。

ここで日本の地方自治体の予算書に注目したい。現行の予算書は、各地方自治体における担当事業部局における目的別支出に応じて提示さ

れる (山本 1991)。この予算は、「地方財政状況調査表」における歳出目的区分に対応する項目として「款」、「項」の議定科目、同調査表で明示を定めている行政科目の「目」からなる (山本 1991)。日本の市町村では、独自の表記の歴史をもつ政令指定都市、あるいは都道府県の指揮監督のもとに置かれていない都市を除き、この「款・項・目」あるいは目の区分としての「節」を加えた予算構造で予算書 (歳出予算書) が統一されている (山本 1991)。

この区分が導入されたのは、地方自治法の昭和38年改正からである (佐藤 2012)。改正前の地方自治法では、財政民主主義の浸透に伴い、予算原則が確立され「予算の編成・提案権は長に委ねつつも予算を議会の議決に係らしめるという予算の議会による統制」(佐藤 2012, p.32)が制度化されていた。その際、首長が提案した予算を議会がどこまで修正できるか、という点が問題とされてきた。この問題に対応すべく、昭和38年改正の際、歳入歳出予算は、款・項によって構成されるもので、目・節は予算説明書である事項別明細書に記載されることが条文として明記された (佐藤 2012)。また、地方自治法第97条第2項において、議会による予算の増額修正は認められるが、首長の予算の提出の権限を侵すことはできないと定めることにより、首長の予算提案権と議会の予算統制権の調整を図っている。そして議会の議決科目である款・項についてもそれを新設する形での予算の増額修正は許されない等、歳入歳出予算の様式を活用して予算をめぐる長と議会の調整や議会による予算統制が行われることとなっている。このように、日本の地方自治体では、議会による予算の統制を明確にするために、予算を款・項・目・節によって編成し、款・項を議会の議決科目、目・節を首長の執行科目としている (佐藤 2012)。

この款・項・目・節の体系は「標準化された政策の入れ子構造」(喜多見 2017, p.307)とも捉えることもできる。款・項・目の構造により、政策目的が数値として記述され、節により、政策の執行方法が数値で記述することが可能となっている (喜多見 2017)。以上のことから、節は、目の再区分として存在しているわけではなく、

款・項・目・節に体系づけられることで、現行の予算の費目分類は、財務情報と政策と事業執行手法の3者を表現した記録（喜多見 2017）となっている。

### 3. COFOGの活用

#### (1) COFOGとは

予算・財政システムの説明責任、および透明性の向上を目的として、本稿が提案する予算科目の分類であるCOFOGとは、Classification of the Functions of Governmentの略称である。これは「一般政府の機能別最終消費支出」と訳される。その意味するところは「一般政府の取引（最終消費支出、中間消費、総固定資本形成、資本及び経常移転に対する支出を含む）を機能又は用途別に分類したもの」（World Bank 2015）である。この一般政府というのは国や地方自治体に代表される行政組織のことを意味する。

COFOGは、国連にて国際的な基準として採択されたものである。COFOGは、国民経済計算システムに含まれる国や地方自治体の支出データを資金の使用目的別に分類するものである。これは、政府支出、すなわち行政組織の活動をその大まかな目的や目標（教育政策、社会保障政策、住宅政策等）に沿って分類する。そのため、分類対象となる行政組織の行政構造や組織構造とは無関係に分類が可能である。この分類により、セクター間の資源配分の分析や、貧困削減のために行った政策に関連した支出を追跡することが可能となる（Jacobs, Hélic & Bouley 2009）。

COFOGに基づく分類は、以下の手順を踏む。まず、機能別に分類した10項目（一般公共サービス、防衛、公共の秩序と安全、経済業務、環境保護、住宅・地域アメニティ、保健、娯楽・文化・宗教、教育、社会保護）に分類する（Ercolano & Romano 2018）。この分類を第1レベル（First Level）とする（OECD 2011）。その後、第2レベル（Second Level）として、第1レベルの各項目を下位の項目に分類する。OECD（2011）によると、第2レベルに相当する項目は、第1レベルの1項目あたり最大 9項目まで増やすことが

でき、全体で69項目存在する（OECD 2011）。さらに、この第2レベルの項目について、各項目を構成する下位項目として第3レベル（Third Level）を設定する。この第3レベルに該当する項目は全部で109個設定されている。以上を整理すると、第3レベルに該当する109個の項目が第2レベルに該当する69項目を構成し、第2レベルに該当するそれらの項目が、第1レベルの10個の項目を構成するという構造<sup>6</sup>になっている。したがって、COFOGは3つの階層からなる（Jeong 2018）分類である。第1レベルは、政府支出の目的の大枠を示し、第2レベル、第3レベルでは、それぞれ第1レベルで示した目的を達成するための手段の詳細な説明となっている（Jeong 2018）。

COFOGデータは、政府支出、特に国レベルでの支出を示すデータとして用いられている。それらの研究の一部を紹介すると、COFOGに基づくトルコ政府支出データとGDPの間のグレンジャー因果性<sup>7</sup>の検証（Arestis, Şen & Kaya 2021）や、COFOGに準拠した政府支出の年次データの分析により、政府支出とGDPとの間の関係が明らかになる（Szarowska 2013）として、その分析により、GDPと総政府支出、GDPと治安維持費、GDPと経済援助費の間の長期的な関係を示した研究（Szarowska 2014）がある。

このように、COFOGデータは国レベルの支出とGDPとの関係を扱った分析・研究で用いられている。これは、COFOGが示す政府支出データとGDPの関係の検証が、先進国において持続的な経済発展を実現する政策の推進に寄与できることが期待されているからである（Jeong, Lee & Kang 2020）。そして、研究成果に基づき、政府支出のポートフォリオが組み直され、教育費や研究開発費の配分を見直されることに繋がり、結果として経済発展の促進に繋がるのが期待される（Jeong Lee & Kang 2020）。既にEU加盟27か国<sup>8</sup>が既にCOFOGデータを公開しており、複数年にわたる基準で支出を判断し、整理する上で大いに有用となるのが期待される（Donath & Miloş 2009）。

このように、COFOGによる国レベルの支出データを用いた研究が行われているものの、課題も指摘されている。例えば、Donath & Miloş

(2009) は、COFOGについて、比較可能性の向上と同時に、政府支出の詳細分析を行う上で、分類されるグループ数が少なすぎる可能性を指摘している。その他、COFOGに基づくデータ、およびそれに基づく分析の課題が指摘されている。例えば、政府支出等の公的支出の分析時には、その支出の効率性と支出によって得られた成果、いわゆるパフォーマンス評価が重要になる (Donath & Miloş 2009)。しかし当該分析に必要なデータが入手困難、あるいは統計的な推定方法が限定されることはよくある (Donath & Miloş 2009)。加えて、一部の国のデータは入手可能であるが、別の国のデータは非公開であること、さらには統計的手法の違いにより国家間での比較ができないことが指摘されている。そのため、国や準国家機関の間で、互いに費目等が異なる場合があるため、政府支出を一般的な分類に基づいて分類したとはいえ、比較する際は、その比較結果はあくまで一般論に過ぎないことに注意する必要がある (Dewan & Ettlinger 2009)。さらに、COFOGによるデータを用いた研究は、国レベルの支出については既に行われているものの、地方自治体レベルについては、その研究はあまり行われていないというのが現状である。

## (2) COFOGによる会計情報の再区分

### ① 利点1: 説明責任の向上

2003年に日本公認会計士協会が公表した「公会計概念フレームワーク」によると、公会計の目的は責任の明確化と意思決定への有用性にある (日本公認会計士協会 2003)。ここでは、国の受託責任として、納税者である国民に対して、徴収した税金等の資源の利用の顛末を国民に対して説明する義務 (山田 2005) が述べられている。地方自治体においても、同様に住民に対する行政責任に基づき会計上の説明責任を果たすための情報開示が求められている (齋藤 2002)。地方自治体と住民との間には、税金等の形で住民から財産の運用と管理についての委託と受託の関係がある (齋藤 2002)。そこで齋藤 (2002) は、「納税者である住民および補助金提供者である国家を委託者とし、行政側を受託者とし、行政側が特定の目的のために使用することを委託

された財産の利用や現状等について住民に説明する義務」(齋藤 2002, p.163) と定義する。

齋藤 (2002) は、行政側への住民参加が「陳情・要請型から、提案・建議型、さらには行政機関と対等のパートナーとしての協働型へと移行」(齋藤 2002, p.165) し成熟する過程において、住民意識の成熟、行政機関による支援体制整備、民主的な組織構成と運営以外に必要な条件 (齋藤 2002) であるとする。一方で、住民に興味関心がなければ、会計情報公開の有用性を論じようとも有効に活用されることを期待することは難しい (山田 2005)。山田 (2005) は、日本とアメリカとの比較から、日本の住民は、納税者として政府に税金を委託するという意識、および納税者としての意識が希薄であり、それゆえに行政組織からの情報の公開の有効性が低下してしまう (山田 2005) ことを指摘している。

ここで課題となるのが、日本の住民の「納税者」としての意識が希薄である (山田 2005) 理由である。この理由について、本稿では、各住民が、自分の払った税金の使われ方が不明であるために、税金委託者 (納税者) としての意識が希薄になっている可能性を指摘したい。日本の地方自治体では、会計上の説明責任を果たすべく、企業会計に準じた財務諸表を公開している。しかし、住民からすれば、これらの財務諸表からでは、自分の住む自治体がどのような政策を行い、その政策を行うにあたってどの程度コストを掛け、どのような成果を出したのかわからないものと考えられる。現行の財務諸表は、企業会計に準拠し、自治体全体の財務状況の提示に留まっている。そして、それらの財務諸表の理解には、会計上の知見を必要とするからである。そこで、本稿では、より細かな分類としてCOFOGによる分類を提示したい。この分類により、各住民に対して、各地方自治体がどのような政策や事業にどの程度税金等を用いたのかを分かりやすく示すことが可能となる。COFOGに基づく分類の利点は、統一された基準に基づいて自治体間の横比較が可能になることと、使用目的ごとに歳出が区分されることにより、住民にとってわかりやすい会計情報となることである。これにより、住民の納税者としての意識

が高まるので、情報公開の有用性を高めることが期待される。

## ② 利点2：透明性の向上

日本の地方自治体の会計制度には、説明責任に加えて、財政・予算の透明性の確保が求められている。財務省財務総合政策研究所（2002）は、Kopits & Craig（1998）を引用し、財政の透明性を「一般国民へ、政府の組織・機能、財政政策の意図、公的部門の会計、財政の見通しが公開（openness）されていること」（財務省財務総合政策研究所 2002、p.1）と定義している。財政が透明性を伴うことで、財政に関する情報は「政府活動に関する、信頼できる、包括的な、時宜にかなった、理解可能な、国際的に比較できる情報を含むもの」（財務省財務総合政策研究所 2002、p.1）となる。さらに、この情報の活用により「選挙民や市場は、政府の財政状態、政府活動の便益とコスト、それらの現在と将来を正しく評価できるようになる」（財務省財務総合政策研究所 2002、p.1）という。

さて、OECD（2000）によると、透明性とは「政策の意図、立案、実施について公開されていることであり、良好な統治（good governance）の最も重要な鍵」（財務省財務総合政策研究所 2002、p.2）を意味する。そのため、財務省財務総合政策研究所（2002）は、予算の透明性を、関係する全情報を適切な時期に体系的な形で公開すること（財務省財務総合政策研究所 2002、p.2）と定義している。また、Blondal（2000）は、財政に関係する情報の体系的な公表が透明性向上の第一要件とする。その上で、公開された情報や報告書の精査・吟味や国民・NPO・メディア等による影響力の行使が重要であるとする（財務省財務総合政策研究所 2002）。

予算・財政の透明性が求められる対象は、「政府の範囲と定義、情報の内容や提供の方法、手続きの方法やそのルール」（財務省財務総合政策研究所 2002、p.6）等広範囲に及んでおり、積極的な推進が求められている（財務省財務総合政策研究所 2002）。しかし、推進にあたって注意すべき点（財務省財務総合政策研究所 2002）もある。すなわち、①透明性そのものは、財政政策

の具体的な方向を示すものではなく、責任ある財政政策のための処方箋でもない。透明性は、あくまで政府活動の状態を明らかにすることにある。そのため、いかなる政府活動、財政活動を行うかは、政府の責任、あるいはそうした政権政党を選挙で選択する国民の責任にある。②透明性向上にはコストがかかる。これは従来の予算・財政システムとは異なる報告書の作成等のために、情報提供システムの構築や人員確保、将来予測の技術的な向上やデータの収集・開発等が必要になるからである。③政府の情報の公開は、方法やタイミングに注意しなければならない。加えて、報告書体系の整備も必要とされる。すなわち、予算・財政の透明性の向上に向けて、予算編成時以外における予算財政情報の開示を進めることや、予算前報告の充実、年次報告及び長期推計報告の作成を進め、諸外国との比較容易性の向上を図る必要がある（財務省財務総合政策研究所 2002）。

## ③ 利点3：住民目線の会計情報

予算や会計情報の透明性を向上させるために、政府によるそれらのデータの公開が促進されることが望まれる（Matheus & Janssen 2020）。データが公開されることで、住民が政府の機能をより詳しく知ることが可能となり、政府の意思決定の改善や説明責任等の効果が期待されるからである（Matheus & Janssen 2020）

正確かつ迅速に提供される予算情報は、政府や住民が予算を正しく選択し、民主的に予算運営が行われ、長期的な財政課題に対処するうえで不可欠である（Reddick, Chatfield & Puron-Cid 2017）。また、会計情報の透明性を高めるという点でも有益なものである（Balaguer-Coll & Brun-Martos 2021）。スペインの地方自治体を対象としたBalaguer-Coll & Brun-Martos（2021）は、政府が公開する財務情報は、野党からの影響に加え、その影響以上に自分たちが納めた税金の使われ方を知りたいという住民からの情報公開要求が高まれば高まるほどその透明性が高まることを指摘している。ただし、その際、公開された会計情報が住民にとってわかりやすいものであるかが重要になってくる（Manes-Rossi 2019; Kim et

al. 2016; Jastrzębska 2020)。Manes-Rossi (2019) は、住民に対して公開される多くの財務報告書が専門的な会計知識を必要とするため理解が難しく、見せている以上に隠していることが多いことを踏まえ、会計情報と会計以外の情報をシンプルでわかりやすい言葉で住民に伝えるために、統合大衆報告 (Integrated Popular Reporting) の開発・利用を提案している。Jastrzębska (2020) も、行政の収入と支出のプロセス、そのプロセスの効果、および経済や社会への影響に住民が興味・関心を抱きづらいことを指摘している。それゆえに、政府として徴収された租税等の入手した資源で何をしようとしているのか、住民が理解できるように設計され、国家予算や予算編成方法に関する専門的な知識を持たない人でも利用可能な住民のための予算を設計する必要があるとする。これにより住民に対してわかりやすい財務情報が提示されるので、自治体会計の透明性を高め、意思決定プロセスへの住民の参加を促進することが期待される (Manes-Rossi 2019; Jastrzębska 2020)。

## 4. 日本での活用にもつれた課題

### (1) 日本での整備状況

日本の予算分類において特徴的である「款・項・目・節」の区分であるが、この区分に基づき、予算を目的別分類、あるいは性質別分類で分類することができる。目的別分類とは、上記の款・項の区分を基準に、各予算を議会費、総務費、民生費、衛生費、労働費、農林水産業費、商工費、土木費、警察費、消防費、教育費等に分類するものである (林 2021)。この分類は、同一基準の分類数字 (コード番号) が付されるため、長期時系列分析に適しているという利点がある。その一方で、上記分類に基づく予算書には「主要経費別に経費が分類、記載されておらず、また主要経費別の歳出総額も記されていない」(三瓶 2020, p.1) という欠点がある。しかし「様々な観点からの分析が可能となるよう各経費にコード番号が付され、各種分類における区分が分かるようになっている」(三瓶 2020、

p.2) ことで、経費区分の分析は可能となっている (三瓶 2020)。

次に性質別分類である。これは、各予算を、上述の節を基準に、人件費、物件費、維持補修費、扶助費、補助費等、普通建設事業費、災害普及事業費、公債費、積立金、投資および出資金、貸付金、繰出金等に分類するものである (林 2021)。この分類は、総務省が決算報告を求める区分である。その意味するところは、当該地方自治体が行政活動として投入した資金量や民間に還流する資金量である (林 2021)。また、地方自治体の財政運営の健全性や弾力性を財務管理の立場から分析する場合には、この分類が有益となる (林 2021)。しかし、この区分と目的別区分、さらに各々の財源内訳<sup>9</sup>からなる支出分析が行われているが、「款」の目的別区分に留まっており、各事業についての具体的な分析には役立ってはいない (山本 1991)。

目的別分類も性質別分類もいずれも「款・項・目・節」の区分に基づいている。この区分は、地方自治法第216条や地方自治法施行規則第15条にその実施が明記<sup>10</sup>されている。しかし、款・項の設定は義務付けられているが、その区分は例示に留まるのに対し、節については、設定のみならず区分も法令で義務付けられている (喜多見 2017)。すなわち、区分項目、特に「款・項」<sup>11</sup>については、地方自治体ごとに柔軟に設定可能なため、統一された項目が使用されていない。そのため、日本では各自治体間で予算の類型化と単位・名称が異なっている。一例として、横浜市と川崎市のように、平成23年度決算において費目に違い<sup>12</sup>が生じている (横浜市 2012; 川崎市 2012)。

予算書上の「款・項・目・節」の構造は、各地方自治体の事業執行内部、すなわち「局 (都道府県の場合)・部・課・係」に従うものとして表示されることが想定されている (山本 1991)。しかし、「款・項・目・節」の構造からでは、各事業の個別性の判別が難しく、特に、必ずしも「目」が事業と事業内容を示す最小単位であるとは言えない状況もある (山本 1991)。また、「款・項・目・節」については議会の審議資料として提出されるものの、議会で審議と議決を必要と

するものは款と項までである。山下（2009）は、日本の地方自治体の会計データ、特に歳出データは人件費や事務費のような形をとるため、それらの費用が、どの程度地域社会の基礎となっているかを把握することはできないことを指摘している。加えて、当該データを生成するに至った政策選択の幅、意思決定の自由度や政治・行政過程を知ることはできず、しかも各自治体間の比較には不十分である（山下 2009）ことも指摘している。

そこで、総務省は、基本計画部会にて日本政府、および地方自治体の支出をCOFOGに基づいて整備することを提案した（総務省 2008）。2008年当時のCOFOGは、10項目の1桁分類 からなるものであったが、当時のOECDは、より詳細な分類を行うべく2桁分類によるデータの収集を行っていた。そこで総務省（2008）では、この2桁からなるCOFOGに基づき、国や地方自治体の支出を分類することで国際間の比較可能性を高め、政策立案に大きく寄与できることを提案していた。加えて、総務省（2008）はデータ分類に際しての具体的な行動も示しており、国の場合は、内閣府が中心として、総務省をはじめとする関係府省等の協力のもと、現行の国民経済計算推計作業で収集しているデータを当該2桁分類に分類することとした。そして、地方自治体の場合、総務省の「地方財政状況調査」の分類と対応が取れる項目の整備、および対応がとれない項目の推計方法について検討し、平成17年基準の改定時を目的にCOFOGの2桁分類による政府支出推計を行うこととした。それは、COFOGの主な利用法として、国や地方自治体の最終消費支出のうち、どの部分が個別に家計に恩恵を与え、どの部分が集合的に家計に恩恵を与えるのかを識別すること（World Bank 2015）を想定していたからである。なお2021年現在、COFOGに基づく国のデータ<sup>13</sup>は整備できている一方で、地方自治体のデータは整備できていない<sup>14</sup>。

## (2) COFOG導入のメリットと課題

COFOGを導入することの大きな利点は、「款・項・目・節」からなる予算分類と比べて、住民にとってわかりやすい会計情報を整備可能であ

るという点にある。住民に対して、わかりやすい会計を整備することは、自治体の透明性を高めることになり、結果として前節で提示した利点を享受することができるものと期待される。

「款・項・目・節」による分類の課題は、その分類によって示される費目が専門用語に基づいており、公開情報といえども、内容を理解するのに相応の知見が求められることにある。Kim et al. (2016) も、現行の行政組織の会計分類システムが、住民に対する情報公開とそれらの情報の活用を念頭に置いて設計されておらず、専門用語で構成されているために、住民にとってその情報を理解することが難しいことを指摘している。そのため、透明性を向上させるために、COFOGを用いて分かりやすい会計情報に加工することが必要だとする（Kim et al. 2016）。具体的には、言葉や数字に加え、様々な図表やグラフ等を用いて、誰にでもわかりやすく提示する（Kim et al. 2016）方法がある。

日本の地方自治体では、従来「款・項・目・節」に基づく予算分類が行われ、予算・決算が報告されてきた。しかし、昨今の社会情勢の変化に伴う新たな課題への対応として、官民協働による対応が普及している。この官民協働がなされるためには、住民が当該地方自治体の取り組み内容等を理解することが求められる。その際に重要となるのが、行政組織による公開情報（Open Government Data）の運用である。そのような取り組みは、既に行われており、イギリスの自治体予算を対象として運営されていた「Youchoose」（川島 2013）、あるいはOpen Knowledge Foundationが主導する「open spending」を代表例として、世界中の公的な財務情報を追跡し、可視化するコミュニティ主導のプロジェクトが始まっている（Kim et al. 2016）。また、Harrison & Sayogo（2014）によると透明性は参加と説明責任に概念的に結びついており、この3つはオープンガバメントに不可欠な要件である。実際、Kim et al. (2016) は、予算データの表示と参加の仕組みを改善したことで、会計情報の透明性の向上のみならず、住民から行政に対する積極的な意見が出るようになり、行政に対する住民参加の意欲が高まったことも論じている。

### (3) オープンデータとしての公会計情報と評価

現在、Open Knowledge Foundation等英国・欧州を中心として、住民レベルによる地方自治体予算の可視化が行われている。この可視化の基盤となっているのがCOFOGである。このCOFOGを用いることで、予算の使い道をよりわかりやすく提示することができ、会計に関する専門的知識が十分でなくとも、行政組織が現在行っている、あるいはこれから行おうとする活動を理解することができる。行政に参加するにあたり、行政の活動を理解することが前提となることは容易に想定できる。このことから、公会計情報をわかりやすく、利用しやすい会計情報とする上でCOFOGは適しているといえよう。

公会計情報をオープンデータ化し、「開かれた行政組織 (Open Government)」を目指すうえで、COFOGによる分類には、地方自治体の説明責任、および情報の透明性を高めることに寄与し、汚職や贈収賄等の不正行為の抑制がなされることが期待される。しかし、データが公開されているが故に、行政の判断や円滑な運営に支障をきたす可能性もある。また、オープンデータとしての会計情報は有用であるとはいえ、そのような会計情報を整備するための業務負担が新たに発生することになる。そのため現行の業務において、業務負担を抑制しつつ、オープンデータとしての公会計情報を整備する方策を今後検討していく必要がある。

また、COFOGによる予算分類と行政評価との連携も期待される。昨今では、統一的な基準（総務省 2019）の導入により、日本の地方自治体における事業別フルコスト用法が得られるようになってきている。しかし、前述の横浜市と川崎市の例にあるように、事務事業の単位は地方自治体ごとに異なっているのが一般的である。そのため、自治体内での行政評価は行えたとしても、自治体間での相互評価という形での行政評価の実施は難しいというのが現状である。そこで、COFOGを導入することで事業単位の統一を行う事ができれば、例えば事業単位当たりのコストなど評価結果の相互比較が可能になることが期待される。加えて、日本の多くの自治体が事務事業レベルで毎年度の予算書上を作成している

ことを踏まえると、予算書上のデータから直接COFOGに変換できるため、比較的作業が容易であり、有用な会計情報となりうることも期待されよう。

## 5. おわりに

地方自治体の予算は、そこに住む住民からの納税等の形で提供された資金と公的に所有されている資産に基づいているため、住民は予算の透明性、参加、説明責任を果たす権利がある。それゆえに、公的予算の透明性は確保される必要があり、そのためには、住民は、公的資金の調達、配分、支出、会計処理の方法に関連するすべての情報が入手可能であり、しかも迅速かつ理解しやすい方法で住民に提供されなければならない (Harrison & Sayogo 2014)。加えて、行政が説明責任を効果的に果たすことで、行政府内の不正や腐敗を抑制し、政府関係者の責任感を高め、設定された政策や事業の目標が達成されない理由について住民の理解を深めることに繋げることができる。そして、最終的には信頼の確立につなげることができる (Harrison & Sayogo 2014)。このように、行政機関の透明性を向上させ、行政への住民参加を促し、住民参加型社会の醸成<sup>15</sup>という点からも住民にとっての「会計情報のわかりやすさ」が重要であることが示唆されている。また、地方自治体側としては、住民とともに会計情報を利用するという理解のもと、会計情報の公開に努める取り組みも求められよう。

会計情報の整備に加え、行政評価制度の導入によっても行政機関の透明性の向上が取り組まれてきた。しかし、評価制度を導入したことで、以前より説明責任が果たされるようになったとはいえない (南島 2013)。説明責任の重要性が、地方自治体側からではなく住民から提唱されたこともあり、地方自治体の現状は「住民からの信頼回復に資するための情報公開」(山崎 2011, p.176)に留まっている。そのため、地方自治体における説明責任を果たす上で、住民に積極的に説明していく姿勢が示された公会計情報の整

備（山崎 2011）が求められる。そこで、本稿で考察したCOFOGを用いることで、現在整備されている会計情報・財務諸表に加えて、住民視点に立つより細かな会計情報を用意することが可能となる。より細かい情報を提示することができるので、住民に対して、自治体運営の全体像を理解しやすく、会計情報の利活用、住民参加の促進に繋がる可能性があると考えられるからである。欧米等での事例で示されているこの可能性が、日本においても同様に見られるのか、さらには、これらの会計情報が、実際の程度活用できるのかについては、紙面の都合上、別稿で扱うものとし、今後の実証分析等により明らかにすべき課題としたい。

## 注記

- 1 内閣による国会への報告は、地方財政法第30条の2の規定に基づいて行われる。
- 2 全国の地方自治体の財務関連情報は、総務省のWebページにて公開されている ([https://www.soumu.go.jp/iken/jokyo\\_chousa\\_shiryo.html](https://www.soumu.go.jp/iken/jokyo_chousa_shiryo.html) 2021年10月31日閲覧)。
- 3 会計情報を用いた政策等の効率化、および財政健全化に向けて行動するという情報利用者としての視点（会田 2013）が十分とはいえない（松尾 2020）ことが指摘されている。
- 4 3つの原則は、具体的には以下の通りである。包括性の原則：予算は、国・地方自治体を含む要請業務を行うすべての行政機関や組織を対象とし、これらの業務の連結された全体像を示すこと、単一性の原則：予算にはすべての国のすべての収入と支出が含まれること、内部一貫性の原則：予算の異なる構成要素間で予算内容が一貫していること。これらの原則は、経常支出と資本（または開発）支出の予算の作成と執行の責任が、財務省のような単一の中央財政機関に集約されていることを意味する。
- 5 COFOGには、IMFによって提唱されたCOFOGとOECDによって提唱されたCOFOGがある。両者の大きな違いとして、IMFはGFS（政府財政統計）ベースのCOFOGを推奨し、OECDはSNA（国民経済計算）ベースのCOFOGを採用しているという点を挙げることができる（Jeong, Lee & Kang 2020）。GFSベースのCOFOGは発生主義に基づくものであり、SNAベースのCOFOGは発生主義と現金主義の両方に基づいている（Jeong 2018）。特にGFSベースのCOFOGであれば、各国間の国際比較が容易にあり、国際基準に基づいた政府支出の体系的なモニタリングを行うことが可能である（Jeong 2018）。
- 6 OECD（2011）では、第1レベルをDivisions、第2レベルをGroups、第3レベルをClassesと称している。
- 7 グレンジャー因果性とは、ある時系列が別の時系列の予測に役立つかどうかを判断するための統計的仮説検定である（Granger 1969）。
- 8 Donath & Miloş（2009）の発刊当時、EU加盟国は27か国であった。
- 9 国庫支出金、都道府県支出金、使用料・手数料、分担金・負担金・寄付金、財産収入、繰入金、諸収入、繰越金、地方債、一般財源等が相当する（山本 1991）
- 10 地方自治法第216条は、歳入歳出予算の区分を示したものである。その条文は「歳入歳出予算は、歳入にあつては、その性質に従つて款に大別し、かつ、各款中においてはこれを項に区分し、歳出にあつては、その目的に従つてこれを款項に区分しなければならない。」である。また、地方自治法施行規則第15条には、「1 歳入歳出予算の款項の区分並びに目及び歳入予算に係る節の区分は、別記のとおりとする。2 歳出予算に係る節の区分は、別記のとおり定めなければならない。」と規定されている。加えて、この施行規則第15条に関連した別記として、歳入歳出予算の款項の区分及び目の区分、歳入予算に係る節の区分、歳出予算に係る節の区分がある。
- 11 下位の項目分類である「節」については、対象項目が別記にて規定されている。
- 12 横浜市では、議会費、総務費、市民費、こども青少年費、健康福祉費、環境創造費、資源循環費、経済観光費、建築費、都市整備費、道路費、港湾費、消防費、教育費、公債費、諸支出金の16費目、川崎市では、議会費、総務費、市民費、こども費、健康福祉費、環境費、経済労働費、建設緑政費、港湾費、まちづくり費、区役所費、消防費、教育費、公債費、諸支出金の15費目であった。
- 13 国のデータについては、内閣府のwebページ「内閣府 一般政府内部門別の機能別支出・機能別最終消費支出に係る参考系列」<https://www.esri.cao.go.jp/jp/>

- sna/data/data\_list/sankou-cofog/index.html (2021年11月12日閲覧) から入手可能である。
- 14 COFOGに基づく地方自治体のデータを整備する取り組みは、民間レベルでは既に行われている。例えば「Where Does My Money Go?」プロジェクト (<http://old.spending.jp/>) において、一部の地方自治体の会計情報についてCOFOGに基づく分類と可視化が行われている。
- 15 Harrison & Sayogo (2014) は、行政組織によって定められた予算（公共予算）の作成と採択に関する意思決定に住民が参加することは、「すべての住民の基本的な権利と責任」と捉えている。
- ### 参考文献
- 会田一雄 (2013) 「公会計情報の活用」、『会計検査研究』、(48) : 1-10
- 東信男 (2020) 「国民経済計算・政府財政統計の課題と展望—一般政府・公的部門における発生主義会計情報の活用に向けて—」、『会計検査研究』、(61) : 67-87
- 井上徹二 (2005) 「「自治体会計の改革」論の吟味」、『横浜経営研究』、26 (1) : 93-105
- 大川裕介 (2020) 「地方自治体の財務報告におけるキャッシュ・フロー情報の重要性」、『公共経営とアカウントビリティ』、1 (1) : 24-36
- 川崎市 (2012) 『川崎市平成23年度一般会計決算』
- 川島宏一 (2013) 「電子自治体と公共イノベーション」、『赤門マネジメント・レビュー』、12 (5) : 415-428
- 喜多見富太郎 (2017) 「自治体における行政活動の定量的把握」、『産大法学』、50 (1) : 305-327
- 齋藤真哉 (2002) 「地方自治体会計への企業会計的手法の導入に関する検討」、『会計検査研究』、(26) : 163-173
- 財務省財務総合政策研究所 (2002) 『我が国の予算・財政システムの透明性—諸外国との比較の観点から—Budget & Fiscal Transparency』
- 佐藤英善 (2012) 「債務負担行為と歳入歳出予算の法的関係」、『自治総研』、38 (1) : 27-43
- 三瓶朋秀 (2020) 「目別分類による予算の論点」、『経済のプリズム』、(186) : 1-32
- 総務省 (2008) 『基本計画部会第2ワーキンググループ報告書』
- 総務省 (2019) 『統一的な基準による地方公会計マニュアル』
- 総務省 (2015) 『地方公共団体における統一的な基準による財務書類の作成予定』
- 南島和久 (2013) 「政策評価とアカウントビリティ：法施行後 10 年の経験から」、『日本評価研究』、13 (2) : 53-67
- 林宜嗣 (2021) 『新・地方財政』、有斐閣
- 日本公認会計士協会 (2003) 『公会計概念フレームワーク—公共部門の財務情報の作成・報告及び予算編成にかかる概念的基礎—』
- 野田遊 (2008) 「行政に対する信頼と市民の参加意向」、『会計検査研究』、(37) : 69-85
- 松尾貴巳 (2020) 「地方公共団体における発生主義に基づく会計制度の導入と予算管理」、『国民経済雑誌』、222 (1) : 17-34
- 山浦久司 (2016) 「地方公会計制度改革の経緯と現状と課題」、『日本会計研究学会特別委員会報告 新しい地方公会計の理論、制度、および活用実践』、9-18
- 山崎敦俊 (2011) 「行政評価とアカウントビリティに資する公会計：地方公共団体の会計システム構築に関する研究視座」、『作大論集』、(1) : 173-192
- 山下茂 (2009) 「地方税財政の総合的国際比較の試み」、『ガバナンス研究』、(6) : 61-125
- 山田康裕 (2005) 「公会計における財務報告の目的とその問題点」、『会計検査研究』、(32) : 23-32
- 山本栄一 (1991) 「地方公共サービスの事業分析」、『経済学論究』、44 (4) : 1-25
- 横浜市 (2012) 『横浜市平成23年度一般会計決算』
- Arestis, P., Şen, H., and Kaya, A. (2021). On the linkage between government expenditure and output : empirics of the Keynesian view versus Wagner's law. *Economic Change and Restructuring*, 54 (2), 265-303.
- Balaguer-Coll, M. T., and Brun-Martos, M. I. (2021). The effects of the political environment on transparency : evidence from Spanish local governments, *Policy Studies*, 42 (2), 152-172.
- Biondi, L., and Bracci, E. (2018). Sustainability, popular and integrated reporting in the public sector : A fad and fashion perspective, *Sustainability*, 10 (9), 1-16.
- Blondal, Jon. (2000) . *Fiscal Transparency*, Report at the Korea/OECD Forum on Public Sector Reform :

- Challenges and Vision for the 21st Century.
- Bryson, J. M., Crosby, B. C., and Bloomberg, L. (2014). Public value governance : Moving beyond traditional public administration and the new public management, *Public Administration Review*, 74 (4), 445-456.
- Caperchione, E. (2000). Trends and Open Issue in Governmental Accounting Systems : Some Elements of Comparison, In Caperchione, E. and Mussari, R. (Eds.), *Comparative Issue in Local Governmental Accounting*. Dordrecht, NL : Kluwer Academic Publishers. 69-85.
- Célérier, L., and Cuenca Botey, L. E. (2015). Participatory budgeting at a community level in Porto Alegre : a Bourdieusian interpretation. *Accounting, Auditing and Accountability Journal*, 28 (5), 739-772.
- Dewan, S., and Ettlinger, M. (2009). Comparing Public Spending and Priorities Across OECD Countries, *Center for American progress*.
- Donath, L., and Miloš, M. (2009). Public sector efficiency according to COFOG classification in the European Union, *Megatrend Review*, 6 (2), 23-40.
- Ercolano, S., and Romano, O. (2018). Spending for the environment : General government expenditure trends in Europe. *Social indicators research*, 138 (3), 1145-1169.
- Fimreite, A. L., and Læg Reid, P. (2009). Reorganizing the welfare state administration, *Public Management Review*, 11 (3), 281-297.
- Harrison, T. M., and Sayogo, D. S. (2014). Transparency, participation, and accountability practices in open government : A comparative study, *Government information quarterly*, 31 (4), 513-525.
- Jacobs, D. F., Héllis, J. L., and Bouley, D. (2009). Budget classification, *Technical Notes and Manuals*, 2009 (006), 1-14.
- Jastrzębska, M. E. (2020). Transparency of information contained in the state budget from the view of utility for the citizen, *Optimum. Economic Studies*, 101 (3), 3-14.
- Jeong, S. H. (2018). New development : A new COFOG for South Korea, *Public Money & Management*, 38 (7), 535-538.
- Jeong, S. H., Lee, Y., and Kang, S. H. (2020). Government spending and sustainable economic growth : based on first-and second-level COFOG data, *Public Money & Management*, 40 (2), 140-148.
- Kim, N. W., Jung, J., Ko, E. Y., Han, S., Lee, C. W., Kim, J., and Kim, J. (2016). Budgetmap : Engaging taxpayers in the issue-driven classification of a government budget. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1028-1039.
- Kopits, M. G. and Craig, M. J. (1998). *Transparency in government operations*, International Monetary Fund.
- Manes-Rossi, F. (2019). New development : alternative reporting formats : a panacea for accountability dilemmas?, *Public Money & Management*, 39 (7), 528-531.
- Manes-Rossi, F., Aversano, N., and Polcini, P. T. (2019). Popular reporting : learning from the US experience. *Journal of Public Budgeting, Accounting & Financial Management*, 32 (1), 92-113.
- Nasi, G. and I. Steccolini. (2008). Implementation of accounting reforms, *Public Management Review*, 10 (2), 175-196.
- OECD. (2000). *OECD Best Practices for Budget Transparency*, PUMA/SBO (2000)6 /Final (Paris).
- OECD. (2011). *Government at a Glance*.
- Reddick, C. G., Chatfield, A. T., and Puron-Cid, G. (2017). Online budget transparency innovation in government : A case study of the US state governments. *Proceedings of the 18th Annual International Conference on Digital Government Research*, 232-241.
- Steccolini, I. (2004). Is the annual report an accountability medium? An empirical investigation into Italian Local Governments, *Financial Accountability and Management*, 20 (3), 327-350.
- Szarowská, I. (2013). Trends of COFOG Government Expenditure in the EU15, *Proceedings of the 10th International Scientific Conference : Economic Policy in the European Union Member Countries : Selected Papers*, Karvina : Silesian University, 280-290.
- World Bank. (2015). *Purchasing Power Parities and the Real Size of World Economies : A Comprehensive Report of the 2011 International Comparison Program*.

## **Promote the Use of Local Government Accounting Information as Open Data open data for municipal accounting information: Focusing on the classification of expense by COFOG**

Yuichi Ubukata  
Waseda University  
y1\_ubukata@aoni.waseda.jp

Toru Satou  
Japan Productivity Center  
toru.sato@jpc-net.jp

Hiroichi Kawashima  
University of Tsukuba  
hkawashima@sk.tsukuba.ac.jp

### **Abstract**

We would like to point out that local governments in Japan focus so much on the development of accounting information that they lack the perspective of information users, especially residents. In Europe and the United States, efforts are being made to visualize local government budgets from the perspective of residents in an easy-to-understand manner. In this case, a classification standard called COFOG is used. The purpose of this paper is to consider the possibility that COFOG can contribute to the creation of accounting information from the perspective of residents. In addition, we discuss the possibility of the improving accountability and transparency of local governments' accounting information by using COFOG. Local governments will be required to develop public accounting information that shows their willingness to explain to residents in the future. Therefore, by using COFOG, it will be possible to prepare detailed accounting information from the perspective of residents. As a result, it will be easier for residents to understand the overall picture of local government management, and there is a possibility of promoting resident participation.

### **Keywords**

Local government, Accounting information, Open data, COFOG



## 【研究論文】

## 「より良い規制」のための評価システムの条件

村上 裕一

北海道大学

yuichim@juris.hokudai.ac.jp

## 要 約

本稿では、「規制の虜」問題の発生を防ぎ、規制を「より良い」ものにしていくための評価システムの条件について検討する。そのために、まず規制監視機関（ROB）に注目し米欧日比較をするのに続き、規制実験評価の方法と留意点を整理する。ROBにありがちな応答性と専門性のジレンマへは、政治的リーダーシップと情報公開で対処する。ROBは、関連する政策共同体の外かつ上に配置し、しかるべき権限と機能の付与、規制監視に係る適正手続の整備、独立性確保による信頼構築に留意する必要がある。本稿では、日本の特区制度も踏まえて望ましい規制評価システムを検討し、「刺激－反応」型モデルを前提とした規制評価が、まずは明確な問いの下、十分に設計された規制実験において、より効果的に実施され得ると論じる。その上で導き出すのは、因果推論と選択バイアスの問題克服、実施・評価活動の客観性確保、実験・評価後の最終決定権者の裁量縮減という条件である。

## キーワード

米国情報・規制問題室（OIRA）、欧州規制精査委員会（RSB）、  
規制監視機関（ROB）、規制インパクト評価（RIA）、規制実験評価

## 1. はじめに：規制行政学と評価研究を繋ぐ

## (1) 「規制の虜」の予防方策

規制は本来、公益を実現するためにあるものだが、様々な理由で歪んだものになってしまうことがある。アクターがそれぞれレント（超過利潤）を獲得するべく直接・間接に規制当局を取り込んで操り、結果的に公益が実現されなくなるという「規制の虜（regulatory capture）」問題が、その最たる例である。

「規制の虜」には、規制に係る便益が特定のアクターに集中するために起こる場合（例：参入規制）と、規制に係る負担が特定のアクター（被

規制者）に集中するために起こる場合（例：安全・環境規制）とがあるが（村上 2022）、特に安全・環境等の社会的規制で起こりやすく、最近問題視もされている後者をcorrosive capture（規制を腐敗させる虜）と呼んだCarpenter & Moss (2014) は、それをどう予防するかというところにまで踏み込んで検討している。

その予防方策は多岐に渡り、そのうちのいくつかは日本でも効果を発揮し得るが（村上2016b）、評価研究との関係では、そのうちの1つ、すなわち規制当局から一定の距離を置いた執政レベル（例：内閣官房、内閣府）から費用対効果等の規制評価を行う（Livermore & Revesz 2014）という

予防方策が、極めて重要である。

## (2) 規制評価を行う重要性

執政レベルから費用対効果等の規制評価を行うことがcorrosive captureを予防し得ると考えられる理由は、次の3つである。

第1に、規制は、政府（行政機構）全体に対するサブ・システム、すなわち、政府の省庁内部の職員のみならずその省庁の活動の対象者集団をも含んだ一種のコミュニティ（「省庁共同体」）の中で、策定・実施される（森田 2022）。例えば、銀行業への新規参入規制は、金融庁をトップに戴き同分野に通じた同省職員や族議員、そして金融業界関係者等から構成されるコミュニティの中で策定・実施される。自動車の安全規制も同じく、国土交通省をトップに戴くコミュニティの中で策定・実施される。この「省庁共同体」では、規制が（公益ではなく）当該共同体の部分利益を反映したものになる恐れがある。そのため、規制当局から「一定の距離を置いた」機関が規制評価を行い、必要に応じて是正を求めるということに意味が生まれる。

第2に、上記のように規制に是正を求める必要が生じた場合に、同レベルの他省庁部局からの是正要求では、それが聞き入れられる保証がない。規制当局が各コミュニティの代弁者として立ち振る舞い続けるためである。そこで、規制当局より「一段高いところに位置付けられた」執政（例：官邸、内閣官房、内閣府）が規制評価を行い、必要に応じて是正勧告・通告をするというのが効果的である。第1点目の関係では、「一段高いところに位置付けられた」執政からの評価だと、より広い視野で公益を捉えることができるという利点もある。ただし、注意を要するのは、執政とて部分利益に囚われてしまうこともあり得るといふ点である。実際には、部分利益の主張者が執政の威を借りて規制を思いのままにすることもあり得ないわけではない（村上 2016b）。

第3に、EBPM（エビデンスに基づく政策立案）の重要性を挙げるまでもなく、規制は何らかの科学的方法により裏付けられた根拠をもって策定・実施されることが望ましい。したがって、

時に合理性を欠くパワー・ポリティクスではなく、費用対効果分析等、ある程度確立された科学的手法により規制の善し悪しを評価・審査するということにも意味がある。

各アクターに多かれ少なかれレント・シーキングの動機がある中で、本来公益を実現すべき規制に関しては、以上の理由から、規制当局とは別の、一段高い執政レベルからの、科学的手法による評価が重要になるのである。

## (3) 本稿における問いと構成

以上のことを踏まえ、本稿では、「規制の虜」問題の発生を防ぎ、規制を「より良い」ものにしていくための評価の仕組みと、それが効果を持ち得る条件について検討する。具体的には、これまで評価研究であまり取り上げられることなかった規制監視機関（ROB：Regulatory Oversight Body）の仕組み（2.）と、ROBにおける権限配分や適正手続等に関する米・欧・日比較（3.）、及び、規制実験評価（Regulatory Experiment Evaluation）についての近年の研究を紹介し、それを評価研究とリンクさせてその利点や課題等について検討した結果を、日本での実践例である特区制度の考察に用いる（4.）。その上で、本稿での検討を踏まえて、日本版規制評価システムが備えるべき条件とともに、実社会への適用で様々な限界や困難を孕む、「刺激－反応」型モデルを念頭に置いた規制評価が、時間や空間をある程度「仕切った」（範囲が特定され、諸条件がコントロールされた）規制実験評価でこそその強みを発揮し得ると指摘する（5.）。

規制評価における指標設定とリサーチ・デザインの重要性を主張するCoglianese（2012）が主に規制開始後の評価に注目するのに対して、本稿は規制評価の実施主体と手法を検討するものであり、その成果は、規制実施の事前・事後を問わず規制の評価一般に適用可能である。

## 2. 規制監視機関（ROB）の仕組み

### (1) ROBの登場とカテゴリー

規制当局から一定の距離を置いて規制評価・

審査を行うROBには、①行政・執政部門で、規制改革府省、省庁横断型ワーキング・グループ、独立組織として置かれるもの（例：米国行政管理予算局（OMB：Office of Management and Budget）の情報・規制問題室（OIRA：Office of Information and Regulatory Affairs））、②立法部門で、議会内の委員会やテクニカル・ミーティングとして置かれるもの（例：EUの立法・執行過程の中に置かれた欧州委の規制精査委員会（RSB：Regulatory Scrutiny Board））、③司法部門で、規制を拒否し得る司法審査として位置付けられるもの（例：最高行政裁判所であるフランス国務院（Conseil d'État）がある。その他にも外部諮問機関、科学アカデミー、メディア、国際機関等がその機能を果たし得るが、ROBの最適な位置付けは、当該国において規制をどこから監視するのが最適かによる。例えば、議会で定める法律が規制の詳細まで規定している国では、ROBを議会付属機関として位置付けることが有効であり得る。

イギリスで1996年3月に旧貿易産業省から内閣府に移管され1999年4月に名称変更した規制インパクト・ユニットは、①の例である。これは政府他部局と協働して規制を公正で効果的なものにする使命を負い、安全・環境規制が経済成長の足枷とならないよう適切なバランスを見極めていた。政府の諮問機関がROBへと発展したフランスやイタリアにおいて、それはそもそも行政裁判所として、政府の行為が民事訴訟に持ち込まれるのを防ぐ役割を果たしてきた（Renaut 2007）。特にフランスでは、君主寄りの判決を下しがちだった司法裁判所に対して、革命後、権力と富を再分配する行政活動に親和的な判断を下したのが国務院で、今それは規制の適法性判断にその専門性を発揮している（Wiener & Alemanno 2010：313）。

## (2) ROBの正当化要件

ROBの権限行使は、応答性と専門性によって正当性を獲得する。

### ① 応答性 (political accountability)

応答性とは、行政官が職務遂行後に政治家に対して説明する責任を指す。確かにそれを果た

すことでROBは正当化されるが、それは、ROBが主権者から直接・間接に委任を受けているということが前提である。例えば、OIRA室長が国民から選ばれた大統領に政治任用されていることにより、規制行政国家を民主的なものにすることができ、そのようなOIRAこそが、やはり国民・議会からの負託を受けた規制当局に対して正当性をもって審査・勧告できる。

したがって、ROBが規制当局に対して実効的な勧告権等を有するには、大統領や総理といった政府中枢に近いことが望ましい（村上 2015；村上 2016a）。世論の支持・信頼を得ているということも重要である（Carpenter & Moss 2014）。

### ② 専門性 (expertise)

ROBには、規制そのものやそのインパクトを正確に評価しきちんと審査する専門性も問われる。それには、規制の内容とインパクトを実態に即して理解することに加え、それを費用対効果等、説得力ある科学的手法を用いて評価できるということも含まれる。ROBにそうした専門性があるのか、それを担える専門人材がいるか、審査の前提として規制に関する情報が不自由なく調達できるか、不当な政治的圧力を受けずに審査できるかといったことが問題になる。

評価の客観性・中立性を確保し信頼を獲得するためには、ROBの独立性が不可欠である。独立性が「政府他部局、政治、活動の相手方への非依存性」（村上 2013）だとすると、その独立性と上記の不自由のない情報調達とはしばしば両立し難い。これには、情報源の多チャンネル化が差し当たりの対処法となる（村上 2009）。

### (3) ジレンマへの対処法

ROBは応答性と専門性を併せ備えていなければならないが、これらに生じる緊張関係・ジレンマにはどう対処するか。

例えば、大統領や総理がROBと異なる選好を持っていたとする。その時、ROBはその専門的な分析結果を彼らに伝え、場合によっては選好を変えるように方向付けるということがあり得る。あるいは、大統領や総理の選好が専門的な分析結果よりも優先されるということを受け入

れた上で、ROBは規制のインパクトやそれが孕むトレードオフの情報を公開し、社会的議論を喚起することもできる。

ROBの規制監視が実効性を獲得するには世論の支持・信頼が必要だが、それが何らかの偏った唱道 (advocacy) と結びついてしまうと、正当性を失いかねない。ROBの説得力を維持・向上するためにはまずその専門的判断が優先されるべきだが、ROBが部分利益に影響され過ぎないようにするためには、ROBの長もしくは職員自身の身分保障を設けておくことも有効である。

### 3. ROBの米・欧・日比較

#### (1) 米国：情報・規制問題室 (OIRA)

米国では、1930年代以降のニュー・ディール期に規制が多く設けられたのを受けて1946年に行政手続法が制定された。1960年から70年代にかけて環境や労働安全衛生等の社会的規制が増え、また経済成長がやや失速してそれへの実務的・理論的関心が高まるのと併行して、ニクソン政権期 (1969年1月～1974年8月) に規制の部局横断的調整が始まり、フォード政権期 (～1977年1月) には規制当局の規則制定プロセスに対する大統領府の関与・統制がさらに強化された。カーター政権期 (～1981年1月) に規制の経済分析を審査したアドホックな規制分析・審査グループに続き、レーガン政権期 (～1989年1月) にOMBにOIRAが設けられ、規制当局による費用対便益分析とOIRAによる規制審査 (regulatory review) が規制制定プロセスの中心に位置付けられた。なお、そこでは、OIRAが評価結果に不満な場合、規制当局に案件を差し戻す仕組みが構築された。続くブッシュ政権期 (～1993年1月) には議会や様々な利益団体からの「押し返し」があったものの、大統領府の権限が確認され、クリントン政権期 (～2001年1月) には、それまでの規制審査プロセスが維持・強化された。ただし、ここではそれまでの費用対便益分析至上主義が修正された。すなわち、規制インパクト評価 (RIA) の視野を広げ、またその量的側面のみならず質的側面をも考慮に入れて、透明性を

高めることで規制の正当性を高めることを目指した。この方針を基本的に継承したブッシュ政権 (～2009年1月) に続くオバマ政権において、規制審査をより実効的かつ軽負担のものにするべく模索されたのが、規制法令施行後もその必要性や効果について随時見直す遡及的審査 (Retrospective Review) と、規制のないことが妥当かを審査する不作為審査 (Inaction Review) であった。

米国特有の党派対立の中でも、大統領府が応答性と専門性をもって規制を監視し一定の判断を下すということについては、ある程度のコンセンサスが存在してきたと言える。規制審査についてはこれまで、そのコスト削減への拘りや法案作成遅延への批判もあったが、最近では「規制の虜」によって歪んだ規制を公益に適ったものに近付けることを期待する声が高まっている (Wiener & Alemanno 2010 : 313-315)。

OIRAは党派性を帯びない50名の専門職員から構成され、室長は大統領により任用される (さらに1986年以降は上院の承認を受けるようになった)。OIRAは評価情報をOMBと大統領に報告し、大統領はそれを受けて規制当局に介入できる。1960～70年代には裁判所に厳しい規制監視を期待する向きもあったが、裁判所は専門性と人員の面で不足があり、大統領が不満を抱くようになって、当初政府業務の合理化を進めていたOIRAに規制監視をも行わせるようになった。ただしOIRAは、議会の立法活動そのものではなくあくまで規制当局の活動に対して権限を行使している。議会内ROBの立ち上げも長年試みられてきたが、OIRAに匹敵する機関の立ち上げにはついに至らなかった (Wiener & Alemanno 2010 : 316-318)。

このように、米国では、議会が法律で規制当局を指導し、大統領が規制当局にRIAを求め、OIRAがそのRIAを審査するという重層的な仕組みになっている。

#### (2) 欧州：規制精査委員会 (現RSB)

欧州では、EU規制の増加を受け、2002年に欧州委員会がRIAを行う規制審査がスタートした。2006年までその評価情報は欧州委内で共有され

るだけだったが、リスボン条約下の「より良い規制イニシアティブ(BRI: Better Regulation Initiative)」を受けて、2006年11月以降、欧州委各局(DG: Directorates-General)の新規制のRIAを、欧州委事務総局(secretariat general)の影響評価会議(IAB: Impact Assessment Board)が監視するようになった。BRIは欧州委でも優先順位が高く、①プロセス全般の透明化、②エビデンスのある立法・政策立案、③市民や利害関係者の巻き込み、を目指す(HP(2022年6月6日閲覧)参照)。その定期的な会合は、事務総局長の下、企業・産業総局、環境総局、雇用・社会問題・機会均等総局、経済・金融総局等、関係4局の代表者から構成され、党派によらない規制審査の基本方針が貫かれた(Wiener & Alemanno 2010: 315)。

EU規制では、まず欧州委が欧州理事会(加盟各国の関係閣僚から構成)と欧州議会に対し新規法案を提案し、DGが欧州委の採択する指令等に沿ってそれを実施する。この中で、IABは欧州委のRIAを監視する。合議体であるIABは、欧州委やDGと対立関係にはないという点で、規制に拒否権を有するOIRAと異なる。IABは規制当局にRIAを求めるのみだが、OIRAは規制当局に何らかの規制措置も促すことができる。IABは規制を強く拒否する権限を持たないが、事務総局長がコメントとともに説得的な報告書をまとめると、DGの規制改善に繋がる可能性が高まる。IABは事務総局と協働することで、適切なRIAを促すことができるのである(Wiener & Alemanno 2010: 318-319, 321-330)。

欧州では欧州委がRIAを自発的に行い、IABがそれを審査する。OIRAが独任制で議会圧力と対峙し、規制提案までできるのに対し、IABはせいぜい立法提案の改善を言うのみで拒否権がなく、プロセスは欧州委内部で完結する。したがってIABの権限は制約されていると言わざるを得ないが、欧州委自体の規制の質向上に向けたソフト・パワーとして無力ではない。

このIABは、2015年7月に「規制精査委員会(RSB: Regulatory Scrutiny Board)」に引き継がれた。IABが各DGの代表者から構成されていたのに対し、RSBの専門家委員(全7名)はフルタイ

ムでDGとは別の枠組みで採用され、それぞれが分野を限定せず評価を審査する。その点で独立性が高められるとともに、審査対象が広げられた(株式会社富士通総研2016: 5-7)。

### (3) 日本：行政評価局と官邸会議

#### ① 事前評価：総務省行政評価局

日本では、2000年前後に強まった規制改革・規制緩和の潮流と、政策評価制度における規制評価の重要性認識に伴い、規制の事前評価が推進されてきた。2004年3月の閣議決定「規制改革・民間開放推進3か年計画」に規制の事前評価が記載され、また、2001年制定の「政策評価法」が一部の研究開発・公共事業・政府開発援助に事前評価を義務付けていたのに対して、その対象外だった規制の新設・改廃に関しても事前評価を実施し国民に結果を公表するべき、との認識がかなり共有された。そうした中で、2004年度には内閣府規制改革・民間開放推進室の実施要領に基づき各府省が規制の事前評価を実施するようになり、その実績が蓄積されて環境が整った結果、規制の事前評価は2007年10月に義務付けられるに至った(財団法人行政管理研究センター 2008: 2-4)。

政策評価法によるこの義務付けに当たって、同法施行令に対象規制の定義が加筆された。もっとも実際には義務付けの対象を巡って論争が巻き起こったようで、基本方針では、義務付けられた規制以外についても積極的かつ自主的に事前評価を行うよう努めることとされた。

2007年6月の閣議決定「規制改革推進のための3か年計画」では、各行政機関が事後見直しにも規制評価を積極的に実施していくべきとされた。当時の英豪両国における規制制定直後の規制評価制度に着目し、事前評価の結果と併せてできるだけ網羅的に規制のメンテナンスが行われるサイクルを確立することが望ましいとされ、これを以後の検討課題とした(財団法人行政管理研究センター 2008: 17-21)。

2007年9月、「規制の政策評価に関する研究会」は最終報告を公表し、これが以後の規制の事前評価の実質的なガイドラインとなった。そこには、理念(規制の事前評価の目的と期待される

効果)、考慮すべき基本事項(評価内容の組み立て、重要性に応じた時間と労力、透明性と検証可能性、不確実性への対処)、政策立案と規制の事前評価(政策立案のプロセスとの統合、広く意見を取り入れる手続)が記されている。

規制の事前評価の結果は原則、規制の新設または改廃が法律による場合は遅くとも法律案の閣議決定までに、政令以下の下位法令による場合は遅くともパブリック・コメントまでに、それぞれ公表される。規制当局から出される全評価書が総務省行政評価局による点検の対象になることから、行政評価局は本稿で言うROBである。点検は①費用と便益との比較分析、②代替案と規制案との比較分析、③レビュー時期の明示等、ガイドラインに記された内容が適切に記載されているかといった観点から実施される。そして、評価書の記載内容に不足があった場合、行政評価局は具体的な改善方策を示して指摘するとともに、その事実を公表して評価書の質の向上に努める。

行政評価局の淵源は旧総理府外局の行政管理庁や総務庁といった行革官庁であり、その点では執政に近く、それに沿った省庁再編の構想もなかったわけではないが(武藤 1998)、現在では他省と同レベルの総務省の一部局に過ぎなくなっている。官僚制の中に位置付けられた分、ある種の専門性は高まったかもしれないが、執政レベルからの評価ではなくなった分、一定程度の応答性を放棄せざるを得なかった。

## ② 政治主導：規制改革推進会議

規制改革推進会議の淵源は、1995年、村山内閣の下で行政改革委員会に設置された規制緩和委員会にまで遡れる。そこではオリックス・グループの宮内義彦CEOが参与(橋本内閣では座長)を務めた。それは1998年1月、橋本内閣で行政改革推進本部・規制緩和委員会(続く小渕内閣と森内閣では規制改革委員会)に再編され、委員長は引き続き宮内義彦が務めた。それは2001年4月、森内閣と第1次小泉内閣で総合規制改革会議に再編され、議長は引き続き宮内義彦が務めた。2004年4月、第2～3次小泉内閣では規制改革・民間開放推進会議となり、2006年以降、

議長は日本郵船の草刈隆郎会長となった。それは2007年1月、第1次安倍内閣～鳩山内閣で規制改革会議となった(2010年3月末まで。議長は草刈隆郎)。2013年1月、第2次安倍内閣では規制改革会議として再出発した(議長は住友商事の岡素之相談役)。2016年9月、同会議の後継組織として発足したのが規制改革推進会議である。

内閣府本府組織令第32条には、同会議の事務として「経済に関する基本的かつ重要な政策に関する施策を推進する観点から、内閣総理大臣の諮問に応じ、経済社会の構造改革を進める上で必要な規制の在り方の改革に関する基本的事項を総合的に調査審議すること」、その「諮問に関連する事項に関し、内閣総理大臣に意見を述べること」が掲げられており、行革というよりもマクロ経済政策が基本理念と言える。大田弘子議長によれば、「1995年から概ね3年ごとに作られてき」ている「規制改革の会議」は、「委員が議論を引っ張っていく」「委員主導の会議」である(第1回会議議事録)。同会議の主要検討課題について掘り下げた審議を行うため、必要に応じ本会議の下にWGが設置される。調査・審議は、その検討課題に関し学識経験者として任命された専門委員の参画を得て進められる(平成28年9月12日規制改革推進会議決定)。

規制改革推進会議では、規制改革、行政手続の簡素化、IT化を一体的に進めるために行政手続部会が設置された。これは政権が毎年改訂する成長戦略に基づくものであり、「内閣官房での調整の下で他の会議体等において議論」するものと整理されている。その事務局は内閣府規制改革推進室が務める。同室長・次長・参事官はいずれも内閣府政策統括官(経済社会システム担当)付等との併任であり、その出自は旧経済企画庁、旧大蔵省、旧行政管理庁、旧総務庁、旧通産省、公取委等である。同室員が各省からの人員に依りながらも内閣府に置かれていることにより、彼らは各省とのリエゾンとなる。政権と距離を詰めて応答性を確保しつつ、外部専門家と連携することで専門性を確保する一種の理想的なROBだが、特定の省庁の影響を強く受けた政権が恣意的な諮問(アジェンダ・セッティング)をすることがあったり、専門性を各省に

依存せざるを得ないことにより同会議の規制改革案が最終的に各省の既得権益に押し負けることがあったりと、それ特有の限界を指摘することもできる（村上2016b）。

#### （4）比較分析と留意点の導出

ROBの比較を踏まえ、「より良い規制」のための評価システムの留意点を整理しておこう。

##### ① ROBに付与する権限と機能

ROBの権限には、例えば、RIAの質の管理や、事後的／補助的な関与（例：事後評価、規制人材育成）といった比較的弱めのものから、望ましくない規制を拒絶したり望ましい規制を推進したりする強い規制監視に及ぶものまである。OIRAは1981年以降、望ましくない規制を当局に差し戻す権限を有しているのに対し、IABはそうした権限は持たないものの、RIA評価書の再提出を求めることでその意思を示すことができる。その意味で、欧州委事務総局長は実質的な拒否権行使も可能だと言える。

OIRAにとって望ましい規制の推進は、2001年以降、プロンプト・レターズによって行われる仕組みになっているが、これは十分に制度化されていない。ついては、科学アカデミーやNGOのようにOIRAの外からのアジェンダ・セッティングや働き掛けが期待される。IABでは、RIAを推進・徹底する趣旨の方が強い。これは、真にRIAが必要かつ有効でありながらそれがきちんと行われてこなかった実態があり、そのギャップを埋めるミッションをIABが負っているということを示している。RIAの定着を図った上で、ROBにどの程度強い権限を与えるかによってその実効性は決まる。

##### ② 規制監視に係る適正手続

米国大統領令には、OIRAが規制当局のRIAを審査する手続や外部接触に関する情報公開、規制当局の反論機会等が定められている。IABに関する2007年のルールでは、メンバー構成や投票、RIA審査のタイミング、議論の公開、専門性・情報源等が規定されている。

OIRAに規制を拒否された所管当局は、副大統領

等、大統領府の上級者に異議申立ができる。裁判所は大統領令ではなくあくまで議会制定法の解釈・適用者であるため、それが定める範囲でしか規制監視機能を果たせない。同様のことは欧州司法裁判所（ECJ）に関しても言えるが、欧州委に対する規制の弊害是正の期待は近年ますます高まっており、ECJはその比例原則に関する判断において、欧州委のRIAを参照するようになってきている。IABの審査に不服のある規制当局は、欧州委に異議申立ができる。OIRAには規制審査の標準処理期間に関するルールがある。

規制の事前評価に働き掛けるのか、政治的対立が先鋭化しやすい規制分野を審査対象に含むのかといったことも、制度設計上の論点になり得る。いずれにしても、審査範囲が広い方がROBの権限を拡大させ、その審査手法の応用可能性も高まるが、いかなる審査手法を用いるのか、それが果たして当該規制分野に馴染むのかという考慮も必要である。

OIRAは社会的規制に主眼を置いてきたが、特に2001年9月の同時多発テロ以降、国家安全保障や対外政策・外交問題に関しても対象に含めるようになってきている。IABも、規制強化のみならず規制緩和案件、さらには経済政策や各種プロジェクトへも視野を広げてきている。米国の金融規制についてはOMBと議会予算室（CBO）が所管しているが、リーマン・ショックを経て、そうした分野の規制監視の必要性も認識されるようになった。審査対象も規制法からガイドライン等に広がってきているが、権力分立の米国でOIRAは議会制定法を審査対象としないことから、連邦議会用のROB創設が求められている。ただし、ROBの濫立への批判はある。

ROBには審査対象の優先順位をつけることも求められる。OIRAでは2003年以降、インパクトが10億ドルを超える規制には確率論的シナリオを備えた評価を行うよう求めている。それに対しIABでは、比例原則を採用し、インパクトの大きい規制にはそれだけ厳格なRIAをするよう求めている。米国では、議会制定法で費用対便益分析を義務付けている分野とそれを禁止する分野を規定している。欧州では規制の功罪両側面を評価することを求めるのみであり、具体的な方

法は指定していないが、2009年のリスボン条約では環境規制に費用対便益分析を求めている。

### ③ 独立性確保による信頼構築

ROBには、規制当局を監視する必要性に鑑みそれらから独立した存在であることが求められるが、現実的には、それらからの適切な距離をどう保つかが問題となる（村上2013）。

IAB委員は、政策決定部局とは独立して活動するものとされたものの実際には規制当局の代表者が委員を務めていたものであり、利益相反の可能性が拭い切れなかった（前述の通り、RSBへの移行で変更）。これに対しては、IABが各局から独立して機能しているという世評が委員のそうした行動をさらに動機付けていたとの指摘もある。IABでは外部からの専門性調達を前提とし、聴聞には当局の代表者が参画したが、OIRAでは規制当局の立会を認めた上で、外部との接触履歴を情報公開の対象としている。

ROBが独立性を突き詰めるあまり孤立無援に陥るのは得策ではない。ついては、多様な情報源によってROB全体として不偏性を確保できるよう、周囲との関係性を積極的に構築していくことが重要だと言える。

## 4. 規制評価の実践：特区制度を例に

本章では、規制評価がどのように実践され、いかなる問題があるかを見るために、日本の特区制度に着目したい。特区制度は、一部地域で規制を緩和してみるという意味において規制の「実験」であり、その結果を評価・解析することによってその規制緩和を全国展開するべきか否かを判断することが想定されている。ここでその実験の評価・解析を行うのがROBであり、それが期待通りに機能するか、その条件とはいかなるものかが問われることになる。

### (1) 規制実験評価の留意点

#### ① 規制実験評価の制度化

Greenstone（2009：114-115）は、社会にとって「より良い規制」を全国展開し、社会にとって悪

弊となる規制をなくしていくために、規制実験とその評価の安定的・永続的なシステム構築を提唱している。そこでは規制の費用対便益分析やインパクト分析が有力な手段となるが（Radaelli *et al.* 2010）、実際には結果ありきのものであったり因果関係のエビデンスに乏しい事前評価に止まっていたりして、その信頼性には疑義も呈されてきた。

そうした中で、医薬品の治験は、米国では30名程度で先行してそれを実施した上で、その結果が良ければ全国展開していくという点、また、専門家を結集した医薬品評価研究センターで徹底的に実験結果の評価・分析を行うという点で、他の規制に应用されるべき特徴を持っている（Greenstone 2009：116）。ただし、この治験でさえ厳格な対照実験は難しく、当該医薬品の有無による効果の違いが十分には明らかにならないという「因果推論の根本的問題」や、実験群と比較群の間で前提条件により結果が左右される可能性を排除できないという「選択バイアスの問題」が付きまとう。もっとも、無作為抽出被験者のグループ間比較を通じたランダム化により、信頼性はかなり向上してきている（Greenstone 2009：117-119）。

規制実験評価では、規制が全国展開される前に、まず一部地域で、少人数の国民に対して実施してみる。その結果を解析して、「規制の失敗」の悪影響を最小限に止めることが期待される。

これを制度化するには、まず評価活動に予算を付け、データを集め、それを広く公表する必要がある。米国では大統領令によりこれがかなり迅速かつ強力にできる。その上で、その予算が高度なサーチ・デザインとデータ分析のできる独立調査研究機関に配分されることが重要である。そのコストは「規制の失敗」がもたらすコストに比べればはるかに小さいにもかかわらず、社会的な賛同は得られにくい。政府としては匿名化したデータと評価結果を情報公開することでデータの矛盾や誤りを抑止し、結果として信頼性を高めたいところである。

#### ② 規制審査委員会の創設

一般論として人間は自らの失敗を認めたがら

ず、自分に都合の良いお手盛り評価をしてしまうということがあるため、自己評価は難しい。そこで、ROBである規制審査委員会（RRB：Regulatory Review Board）には、関連研究をレビューし、必要に応じて追加的に予算を要求・獲得し、費用対便益分析の結果を踏まえ規制を直ちに止めるのか全国展開するののかについて判断する権限を付与する。米国食品医薬品局が治験結果を踏まえ医薬品の流通を停止できるように、規制法令にRRBの権限が書き込まれることが望ましいが、その場合も、議会や大統領・総理の権限との関係には注意する必要がある。RRB委員長は政治任用としつつ、RRB委員は、評価結果を批判的に審査する専門性を有するとともに、規制に偏った利害のない人物である必要がある。Greenstone（2009）の提案は、特定のイデオロギーに染まっておらず、また官僚制組織の悪弊をなくした不偏不党な合議体であるべきというものだが、ここでは、委員個々人の党派性を問うよりも、委員会全体でイデオロギーに偏りが無いかが問われるべきと思われる。他の例として、米国国立衛生研究所や国立科学財団で研究費配分を行う独立審査委員会がある。

同様の役割はOIRAにも期待されているが、OIRAには、①厳格な費用対便益分析が全規制に必須ではない、②ほとんどが「理論」的見積による事前評価であり、規制「実験」評価ではない、③室長が大統領による政治任用であり、党派性から完全に無縁ではない、といった問題点がある。OIRAをRRBに近似させていくことが望ましいが、規制当局との緊張関係には注意を要する（Mitchell *et al.* 2010）。

### ③ 自動停止・拡張と倫理の規程

RRBが規制実験評価を踏まえ存廃の判断を下す期日を規制法令で定め、判断がないままその期日を過ぎた場合、規制は自動的に停止される。逆に、規制実験評価で望ましい結果が出た場合は、それを既存の規制体系に組み込み全国展開する方向へ持っていく。ここで「自動的な」停止・拡張を強調するのは、その決定が可能な限り規制実験評価の結果に基づくべきと考えるからである。

規制実験評価は一定数の国民の犠牲の上になり立つものなので、倫理的な問題を孕む。これへの対応として、規制実験評価が一定の配慮を伴いながら実施されることを担保するために、①それが国民にとって安全であることを確認の上、②そのための倫理委員会を創設することを定めた研究倫理規程を設けることが必要である。倫理委は政府内規制（政府の自己規制）になるため、対他者の規制にない独特の難しさがある（Lodge *et al.* 2010）。

## (2) 評価研究とのリンク

### ① 政策評価の技法と現実的対応

政策評価の技法は（a）政策を把握する局面、（b）評価基準、（c）評価の対象レベルの3側面から捉えられる（田辺1998：293-295）。

（a）には①費用、②作業、③事業、④効果の4つがあり、（b）には①有効性、②効率性、③公平性がある。費用対便益分析は（b）②に分類されるが、規制実験評価の基準は（b）①や（b）③もあり得る。（b）①は、政策が所定の効果を上げたのか否かが評価対象となるが、本来は行政資源の変換それ自体の質も考慮に入れられるべきである。（b）③では、当初予定していた対象集団に政策の効果が届いているかどうか、費用と効果の配分に各集団の間で著しいばらつきや偏りが無いかといったことが問題となる。

（c）には、下位から順に①個別プロジェクト、②プログラム、③政策の3つがある。定義上、（c）①から②、③へとレベルが上がるにつれてより多くの目標が含まれるようになることから、評価対象の規制が広範囲なものになればなるほど、単純化した「刺激－反応」型モデルから乖離していくことになる。Koenig-Archibugi（2010）は規制が国際化し複雑化する中で評価手法をさらに洗練させていくべきと主張するが、現実的な規制評価とはいかなるものか。

### ② 「刺激－反応」型モデルを巡って

規制実験評価の意義は、規制のインプットとアウトプット、アウトカムとの対応関係をモデル化して捉えることで、「刺激」と「反応」の間の副次的要因をかなり除くことができ、したがっ

てそれらが容易に関係付けられるようになるということである。規制実験評価はそれを可能にする一手法だとも言え、ゆえに、当該規制の存廃判断の重要なエビデンスを提供し得る。指標を金銭的価値に置き換えることによって、異なる分野・環境下でも比較が可能になる。

田辺（1998：297-299）は、この「刺激-反応」型モデルの問題点として、①それらの間をブラック・ボックス化する結果、政策実施に関わる文脈的な知識の蓄積が行われない、②「刺激」と「反応」とを1対1に対応する単位にまで政策を分解するため、大きな視野で政策を捉えることができない、③政策によって枠づけられる社会・経済の構造とそれによって規定される諸アクターの行動は時間が経って政策が作用すれば変わるため、ある一時のモデルからの推計で得られたパラメータを用いて編み出される最適の政策は、また別の場面では最適ではありえない場合がある、という点を指摘する。

これに対し規制実験評価は、時間・空間を「仕切る」ことでこの困難をある程度克服できれば、「より良い規制」に資することが期待される。

### ③ 規制実験としての特区制度

特区制度は、現実には、時限的・地域的例外をそこに押しとどめようとする勢力とその「特別扱い」の恩恵に浴そうという勢力、そしてそれを「蟻の一穴」として全国展開しようとする勢力の妥協の産物として、概ね運用されている。

しかしながら、それは規制実験と捉えることもできる。鈴木（2004：145-146）は、構造改革特区（2002年～）の特徴として、①減税措置や補助金等の財政措置を伴わず、純粋に規制緩和だけを行うものである、②特区における規制緩和が有効に機能すれば、速やかに全国的な規制緩和を実行することが予定されている、③特区における規制緩和はオープンな政策評価が行われ、その評価に基づいて全国展開するか否かが決定される、という点を指摘している。

国家戦略特区（2015年～）も、総理や内閣の主導で国の成長戦略を実現するべく、大胆な規制改革を実行するために創設され、官邸HP（2022年6月6日閲覧）によれば実際に全国展開に繋が

った事例もある（例：養父市の「古民家等の歴史的建築物に関する旅館業法の適用除外」、福岡市の「エリアマネジメントの民間開放（道路の占用基準の緩和）」、新潟市の「農業生産法人6次産業化推進等のための要件緩和」）。

国家戦略特区について、2015年度末までの2年強の間の集中取組期間終了後に公表された初回定期評価は、そのスピーディな進捗を指摘している。しかしながら、高坂（2016）によれば、実際には着手時期が計画後期に後ろ倒しで設定されていたことから、進捗はそれほど早くなく、経済・社会に多大なインパクトを与えた規制特例措置も見受けられない。評価に関しては、①定量評価が行われず定性評価（計画通りの進捗か、着手済みかの2点）のみである、②活用事例の有無のみが評価され、数値目標に基づく成果指標のチェックが行われていない、③評価主体は特区ごとの運営当事者であって、第三者評価ではない、といった問題点もある。

確かに高々数年の規制特例措置後の評価であり、その社会的影響を定量的に測定するのは難しい。しかし、4. (1)で整理した留意点①～③に照らすと、下記の問題点を追加的に指摘せざるを得ない。

第1に、「因果推論の根本的問題」や「選択バイアス」(①)を克服するための配慮や工夫がほとんど見られない。理論的には、同規模・同条件の複数地域・グループ間での対照実験が望ましい。現実には容易でないものの、類似の地域・グループ間で同時並行的に措置し、その評価結果を比べたり対象地域における措置の前後で社会的影響を比べたりするのが有効なはずである。

第2に、各特区に設けられた区域会議が評価を実施していたり、実際の評価が主に事業の進捗状況について行われていたりすることには、評価理論(②)上の問題がある。そもそも地域のニーズを(も)汲んで指定された特区の効果を評価する必要があるのだとすると、その客観性を確保する仕掛けが、適切な指標設定と共に必要である。

第3に、自動停止・拡張の提言(③)を念頭に置くと、特区制度上、総理やその任用による諮問会議が評価結果を基に行う存廃判断に大きな

裁量が残されていることは、評価結果を規制改革に直に反映していくための仕組みとして脆弱である。費用対便益分析の結果を根拠に存廃判断をするかどうか自体には議論のあり得るところだが、その判断に裁量の余地を残すことは、様々な政治的介入・配慮を招くのみならず、無責任な決定を誘発する恐れがある。

## 5. おわりに：「より良い規制」のために

本稿では、「より良い規制」の策定・実施のためにあるべき評価の仕組みとはいかなるものかという問いを出発点とし、ROBや規制実験評価に関する理論的検討と、特区制度における評価活動に関する実践的検討を行った。

各国のROBについては、応答性と専門性のジレンマがそれぞれに見て取れたが、その解消法として、ROBの専門性を基盤として説明責任を果たしていくことや、政治決定（応答性）を尊重した上で科学的評価の情報を公開し社会的議論を喚起することを挙げた。歴とした規制評価が未確立な日本では、ROBについてのしかるべき権限配分と適正手続、そして独立性確保による信頼構築に留意した制度設計により、「より良い規制」のための評価システムの整備が求められる。「刺激－反応」型モデルから見た評価の限界をかなり克服し得る、「仕切られた」時間・空間における規制実験評価の例として、本稿では日本の特区制度を取り上げた。その実践を参考に指摘したのは、①因果推論と選択バイアスの問題克服、②実施・評価活動の客観性確保、③実験・評価後の最終決定権者の裁量縮減という、「より良い規制」のための評価システムの条件である。

本稿の指摘を巡っては、規制実験評価への社会的な理解と寛容が1つのポイントになる。評価理論との関係を踏まえた本研究を通して、ROBの存在意義と規制実験評価のさらなる可能性が明らかになった。ランダム化された定量的評価がさしあたり最善の規制評価手法だとすると（Coglianese 2012）、評価の制度や手法を改善していくために、まずは本稿の規制実験評価をモデ

ルとすることが有効かつ現実的ではないか。我が国で2018年に始まった規制のサンドボックス制度（官邸HP（2022年6月6日閲覧）：新技術の実用化が現行規制との関係で困難な場合に、その社会実装に向け、事業者の申請に基づき規制官庁の認定を受けた実証を行い、それにより得られたデータを用いて規制の見直しに繋げていく制度）は、その実績が蓄積されてくれば、やはり本稿の趣旨に沿った検証の対象となり得る。

## 謝辞

本稿は、2020～2022年度 科学研究費補助金「マルチレベル・ガバナンスの研究：人や組織のdiscretionによる事例分析」（代表：筆者）の成果の一部をまとめたものである。本稿は、日本評価学会の2015・2016・2017年度全国大会で口頭発表の機会を得た際の予稿を基にしている。各自由論題で司会・討論者を務められた和田義郎先生、上野宏先生、佐藤由利子先生、及び、査読者の先生方等、本研究にご指導くださった皆様に深く御礼申し上げる。

## 参考文献

- 株式会社富士通総研（2016）『欧州連合（EU）における規制の政策評価に関する調査研究報告書（平成28年12月）』
- 鈴木亘（2004）「構造改革特区をどのように評価すべきか：プログラム政策評価の計量手法からの考察」、『会計検査研究』、30：145-157
- 高坂晶子（2016）「国家戦略特区の第1回評価と今後の課題」、『日本総研Research Focus』
- 田辺国昭（1998）「政策評価」、森田朗編『行政学の基礎』、岩波書店、284-301
- 武藤桂一（1998）「日本の行政改革における内閣機能の強化」、片岡寛光編『国別行政改革事情』、早稲田大学出版会、291-326
- 村上裕一（2009）「官民協働による社会管理：自動車安全のための技術基準策定プロセスを素材として」、『国家学会雑誌』、122（9-10）：1266-1330
- 村上裕一（2013）「行政の組織や活動の「独立性」について」、『社会技術研究論文集』、10：117-127

- 村上裕一 (2015) 「「司令塔機能強化」のデジャ・ヴ  
ユ：我が国の科学技術政策推進体制の整備を例に」、  
北海道大学公共政策大学院『年報 公共政策学』、9：  
143-168
- 村上裕一 (2016a) 「政治と公共政策：科学技術の省庁  
再編をめぐる」、北海道大学公共政策学研究セン  
ター監修・西村淳編『公共政策学の将来：理論と実  
践の架橋をめざして』、北海道大学出版会、99-130
- 村上裕一 (2016b) 「いわゆるCorrosive Captureとその予  
防方策」、北海道大学公共政策大学院『年報 公共政  
策学』、10：141-165
- 村上裕一 (2022) 「国際海運の環境規制の特徴」、北海  
道大学大学院法学研究科『北大法学論集』、72 (6)：  
1-29
- 森田朗 (2022) 『新版：現代の行政〔第2版〕』、第一法  
規
- 財団法人行政管理研究センター (2008) 『規制の事前  
評価ハンドブック：より良い規制に向けて』
- Carpenter, D. & Moss, D.A. (2014). Introduction. In  
Carpenter, D. & Moss, D.A. (eds.), *Preventing  
Regulatory Capture: Special Interest Influence and How  
to Limit It*. New York: Cambridge University Press, 1-22.
- Coglianesse, C. (2012). *Measuring Regulatory Performance:  
Evaluating the Impact of Regulation and Regulatory  
Policy. Expert Paper No. 1, August 2012.*
- Greenstone, M. (2009). Toward a Culture of Persistent  
Regulatory Experimentation and Evaluation. In Moss, D.,  
& Cisternino, J. (eds.), *New Perspectives on Regulation*.  
The Tobin Project.
- Koenig-Archibugi, M. (2010). Global Regulation. In  
Baldwin, R., Cave, M., & Lodge, M. (eds.), *The Oxford  
Handbook of Regulation*. Oxford University Press.
- Livermore, M.A., & Revesz, R.L. (2014). Can Executive  
Review Help Prevent Capture? In Carpenter, D. & Moss,  
D.A. (eds.), *Preventing Regulatory Capture: Special  
Interest Influence and How to Limit It*. New York:  
Cambridge University Press, 420-450.
- Lodge, M., & Hood, C. (2010). Regulation Inside  
Government: Retro - Theory Vindicated or Outdated?  
In Baldwin, R., Cave, M., & Lodge, M. (eds.), *The  
Oxford Handbook of Regulation*. Oxford University  
Press.
- Mitchell, C., & Woodman, B. (2010). Regulation and  
Sustainable Energy Systems. In Baldwin, R., Cave, M.,  
& Lodge, M. (eds.), *The Oxford Handbook of  
Regulation*. Oxford University Press.
- Radaelli, C.M., & Francesco. de-F. (2010). Regulatory  
Impact Assessment. In Baldwin, R., Cave, M., & Lodge,  
M. (eds.), *The Oxford Handbook of Regulation*. Oxford  
University Press.
- Renaut, M.-H. (2007). *Histoire du Droit Administratif*.  
Paris.
- Wiener, J.B., & Alemanno, A. (2010). Comparing  
regulatory oversight bodies across the Atlantic: the Office  
of Information and Regulatory Affairs in the US and the  
Impact Assessment Board in the EU. In Rose-Ackerman,  
S., & Landreth, P.L. (eds.), *Comparative Administrative  
Law*. Cheltenham, UK. Northampton, MA, USA: Edward  
Elgar, 309-335.

(2022.6.11 受理)

## **Institutionalizing Systematic Evaluation to Ensure “Better Regulation” - Focusing on Regulatory Oversight Body and Experiment Evaluations -**

Yuichi Murakami

Hokkaido University  
yuichim@juris.hokudai.ac.jp

### **Abstract**

To prevent “regulatory capture,” this article discusses a systematic evaluation to ensure “better regulation” and a way to institutionalize it, focusing on (1) the Regulatory Oversight Body (ROB), its dilemma of political accountability and expertise, its comparison with the US, EU, and Japan, and (2) regulatory experiment evaluations. ROB should be placed above and out of the related policy community to ensure meaningful oversight and objective evaluation, and to eventually prevent the regulation from being captured by partial interests. Then, (3) a desirable regulatory evaluation system is envisioned, according to a review of the Special Zone System of Japan. The “stimulus-response” modeled regulatory evaluation can be properly conducted in well-organized regulatory experiments. The main points consist in overcoming the difficulties in causal inference and selection bias, rendering implementation and evaluation activities more objective, and reducing the final decision maker’s discretion after obtaining the evaluation result.

### **Keywords**

Office of Information and Regulatory Affairs (OIRA) , Regulatory Scrutiny Board (RSB) ,  
Regulatory Oversight Body (ROB) , Regulatory Impact Assessment (RIA) , regulatory experiment evaluation



## 【研究ノート】

業績測定による評価における指標の質改善のための  
帰納的情報の必要性

池田 葉月

京都市立大学

ikedahazuki@kpu.ac.jp

## 要 約

業績測定は比較的簡易な評価手法であり、業績測定を中心的な手法とする評価制度は、公共部門において普及している。業績測定は事前に設定した指標を用いて測定した実績値と目標値を比較して評価する手法であることから、指標の質が重要である。そのため、適切な指標を設定するとともに、不適切な指標は見直す必要がある。指標の質を改善する方法には様々なものがあるが、その1つとして、見直す前の指標と見直しの理由、見直した後の指標などの情報を整理し、そこから学習することが挙げられる。しかし、指標の見直しの実施状況や見直しの理由に関する情報は少ないことを本稿では明らかにする。また、指標の見直しに関する情報などの帰納的な情報が不足していることを明らかにし、それらが指標の改善に役立つことを述べる。

## キーワード

業績測定、指標の設定と見直し、自己評価の質向上

## 1. 本稿の背景と問題意識

業績測定は、事前に設定した指標を用いて実績値を測定し、目標の達成状況から評価する評価手法である。他の評価手法と比較すると専門性が低く、得られる情報の正確性や客観性は劣る。一方で、高度な専門性を有しない評価者でも使うことができ、多様な分野に適用可能であるため汎用性が高い。そのため、自己評価の手法として公共部門で広く用いられており、業績測定を中心的な手法として用いる評価が制度として導入されている場合が多い。

評価手法の特性上、業績測定による評価において指標は重要な役割を果たしており、評価結

果の正確性や客観性にも影響する。しかし、指標については課題が指摘されることが多い。

また、業績測定を用いる評価制度では、指標の重要性が高いことから、指標そのものだけでなく、指標に関する情報を整理することも評価制度の適切な運用という観点から重要性が高いと言える。つまり、設定していた指標がどのようなもので、その指標にはどのような問題があり、どのように見直したのか、見直した結果どのような指標をなぜ設定したのかという情報が揃っていないならば、必要な見直しだったのか、適切に見直されたのかを判断することができない。これらの情報が整理されていなければ、指標の見直しが実施されていたとしても、毎年度

の評価表をたどっていかなければ見直されたのかどうか、またどのように見直されたのかわからない。このような状態では、評価制度が適切に運用され、その目的を果たしているとは言えない。つまり、形式的には評価を実施していても、不適切な指標から得た情報では、評価対象の政策について必要性や有効性、費用対効果などを十分に評価することはできない。そのような状態では、評価を通じた政策の改善にはつながらず、アカウンタビリティを果たしているとも言えない。また、どのような場合に指標を見直すべきなのか、どのような見直しの方法があるのかなどの情報も蓄積されないため、適切な指標の設定は継続的な課題となっていくと考えられる。

業績測定による評価における指標の質を改善する方法の1つとして、指標の見直しに関する情報を整理し、そこから学習することが考えられるが、本稿では、実際には十分に整理されていないことを明らかにする。また、指標の見直しに関する情報を整理することで、指標の質改善のための既存の取組みから得られる演繹的な情報に帰納的な情報を加えることができ、指標の質改善に役立つと考えられることを述べる。

なお本稿では、演繹的な情報とは指標に関する一般的な説明を意味している。例えば指標の設定方法に関する一般的な手順や考え方、指標の設定において注意すべき点、望ましい指標に見られる共通の特徴などが該当する。また、帰納的な情報とは、改善すべき指標から得られる具体的な情報を意味している。例えば改善が必要な指標とはどのようなものなのか、どのような点に改善が必要なのか、実際にどのように改善できるのか、あるいは改善されたのかなどが該当する。

## 2. 業績測定による評価における指標の重要性

2では、業績測定による評価における指標について、どのようなものであるべきであるとされているのかということを述べる。

適切な指標は、ロジックモデルを明確化し、それを踏まえて測定すべきものを特定することで設定できる（上野・上野 2008、pp.65-67；小野 2020、pp.181-182；佐藤 2021、pp.28-30；米原 2020、pp.150-151）。そのため評価者は、その指標によって何を測定するのか、設定された指標から評価に必要な情報を十分に得られるかどうかを常に意識しながら指標を設定する必要がある（米原 2020、pp.151-152）。また、偏りや漏れをできるだけ少なくし、質・量ともに十分なデータを収集できる指標を設定することで得られる情報の信頼性も高まる（小野 2020、pp.182-183）。また、単一の指標で測定すべきものを網羅することは困難である場合が多いため、複数設定することが望ましい。しかし、考えられるものを多数羅列するだけでは無駄が発生するとともに分析も困難になる（米原 2020、p.151）。

指標を設定する方法や手順については、様々な文献で示されているが、公共政策によって最終的に実現したい状態を明確化した上で、測定すべきアウトプットとアウトカムを特定し、測定方法も考慮して設定するというのが大まかな流れである。業績測定による評価の手順を詳細に説明しているものとしては上野・上野（2008）やハトリ（2004）などが挙げられる。上野・上野（2008）では、業績測定による評価の過程が評価結果の利用の段階まで含めて10段階に分けて説明されているが、そのうちの4段階目までは指標の設定に関することである（上野・上野 2008、pp.65-71）。ハトリ（2004）は全15章のうち6章は指標の設定に関することであり、指標を設定する前段階として考えるべきこと、アウトプット指標とアウトカム指標の考え方、データの測定方法などが説明されている。また、業績や評価に関する情報を利用して成果志向の意思決定を行うという業績予算の観点からは、予算は目標を達成するためのツールであり、両者の関連付けに留意することの必要性が指摘されている（稲沢 2021、p.40）。

### 3. 業績測定による評価における指標の現状

3では、業績測定を中心的な手法として用いる評価制度における指標の現状を述べる。主要な評価制度として地方自治体の評価と府省の目標管理型評価を取り上げる。

総務省の調査では、地方自治体では指標の設定について課題を認識している割合が高い。評価制度を導入している1,099団体のうち78.5%が課題の1つとして挙げており、主要な課題の1つとなっている（総務省自治行政局市町村行政経営支援室 2017、p.7）。具体的な課題としては、指標が設定されていない、アウトプット指標とアウトカム指標が混同されていることや、不適切なアウトカム指標が目立つこと、目標値の水準が低すぎるまたは高すぎる、事業の内容が十分に具体化されていないなどの理由から達成できるのかどうか不明確であることなどが指摘されている（佐藤 2021、pp.8-9；内閣府地方創生推進室 2015、pp.1-5；内閣府地方創生推進事務局 2020、p.25）。

また、反復的なタスクや創造的なミッションなどの分野には積極的に指標が設定されてきたが、予防や防止などの分野への指標設定には消極的であり、指標自体が設定されていない場合もあるという偏りがあることも指摘されている（坪井 2021、pp.32-33）。さらに、地方自治体における評価制度の実態に関する調査結果の経年変化を見ると、多くの地方自治体で指標が設定されており、必要性も認識されているが、指標の設定は継続的に課題として認識されている（総務省自治行政局行政経営支援室 2011；総務省自治行政局行政経営支援室 2014；総務省自治行政局行政経営支援室 2017；早川・金崎・北山 2021、pp.746-748；三菱UFJリサーチ&コンサルティング 2021、pp.33-35・pp.39-40）。

府省の評価制度における目標管理型政策評価については、政策評価審議会政策評価制度部会がまとめている各年度の「目標管理型の政策評価に係る評価書の検証結果」においても課題の1つとして指摘されている。また、政策評価審議会が2021年3月にまとめた行政における評価のあ

るべき姿の1つとして「しなやかな評価」がある。これは、評価者が時間の経過や社会経済の変化に柔軟に対応して評価しているという状態を意味している。「しなやかな評価」は、指標に関することだけを述べているものではなく、施策の特性に応じて適切な評価手法を選択するとともに、どの程度詳細な評価を実施するかについてもメリハリをつけるべきだというものである。そのために、ロジックモデルなどを活用することによって適切なポイントに適切な指標を設定することが重要であるとされている（政策評価審議会 2021、p.8）。

### 4. 指標の見直しの現状

4では、業績測定を中心的な手法として用いている府省の目標管理型政策評価と京都市の政策評価制度を対象として、指標の見直しはどの程度行われているのかについての調査結果を示す。また、どのような場合にどのように指標が見直されているのかについての調査結果を示す。

#### (1) 調査のデザイン

地方自治体の例としては京都市の政策評価制度を取り上げる。評価制度を導入している地方自治体は多数あるが、指標の見直しの状況を整理して公表している地方自治体は少ない。そのような中で、京都市は指標の見直しの状況を整理して公表していることに加え、業績測定を深めて精緻化しているとされており（窪田 2014、pp.15-18）、先進的な例として扱われることも多い。そのため、本稿の調査対象とする。

府省については、目標管理型政策評価を対象とする。目標管理型政策評価は、各府省の主要な施策を対象として実施されている評価であり、評価手法としては業績測定が用いられている。2016年度から2020年度の間で、事前分析表において測定指標の見直しを行った数が多い上位8つの行政機関である内閣府、金融庁、消費者庁、総務省、経済産業省、原子力規制委員会、文部科学省、厚生労働省について、指標が見直された理由を整理する。

なお、京都市と府省を比較する際には、両者のデータが共通して得られる期間である2017年度から2019年度を対象とする。指標見直しの件数と理由を整理する上で参照したものは、京都市の政策評価委員会で配布されている資料と、各年度の「政策評価等の実施状況及びこれらの結果の政策への反映状況に関する報告」、「各行政機関における政策評価の結果及びこれらの政策への反映状況（個表）」である。

## (2) 京都市の政策評価制度

京都市は、「京都市行政活動及び外郭団体の経営の評価に関する条例」に基づいて総合的かつ体系的な行政評価等の仕組みを構築しており（第1条）、その1つとして政策評価制度を運用している（第7条）。

政策評価制度は総合計画で規定されている政策体系に基づき、市が目指すべき基本的方向である政策（27項目）と、政策を具体化した施策（114項目）を評価対象として2004年度から運用されている。業績測定を中心的な手法とする評価制度であり、客観指標評価と市民生活実感評価の2つの評価結果を総合して評価している。本稿の対象は客観指標評価であり、これは各政策・

施策の状況を統計的な数値等により表現した客観指標を設定し、その目標達成度から評価するものである。また、制度の公正な運用と向上を図るために外部機関として京都市政策評価委員会を設置し、客観指標の選定や評価制度の運用などについて意見を求めている。

京都市では、適切な指標の設定と見直しのために「客観指標の設定マニュアル」を策定している。このマニュアルでは、指標や目標値を設定する必要性、ロジックモデルを用いた指標設定の考え方、指標設定の際に確認すべき点が事例とともに解説されている。

このような評価制度における指標の見直しについては表2のようになっている。京都市では、見直しの理由は定型化されており、見直しの主な理由は「新たに策定・改定した分野別計画の目標に合わせるため」、「より適切に評価するため」となっている。そのいずれかに分類した上で、見直し前の指標と見直し後の指標、さらに詳細な見直しの理由をまとめて公表している。これらの情報からは、以前の指標のどのような点になぜ問題があり、どのように改善した結果、どのような指標になったのかがわかり、帰納的な情報であると言える。

表1 指標が見直された政策と施策の数

見直しの理由	2015年度	2016年度	2017年度		2018年度		2019年度		合計
			政策	施策	政策	施策	政策	施策	
総合計画の実施計画を更新したため	×	×	4	2	0	0	0	0	6
新たに策定・改訂した分野別計画の目標に合わせるため	○	×	3	8	0	0	7	7	25
より適切に評価するため	○	○	4	4	2	14	2	5	31
目標達成に伴い、指標の目標値を見直した	○	○	1	5	0	1	0	0	7
合計	—	—	12	19	2	15	9	12	69

単位：件

(注1) 2015年度と2016年度は、見直した理由は公表されているが、件数は公表されていない。

(注2) ○は該当する政策または施策があること、×は該当する政策または施策がないことを意味している。

(注3) 異なる施策で同じ指標を用いている場合と、政策と施策で同じ指標を用いている場合もそれぞれにカウントされている。

(出所) 各年度の政策評価委員会で配布された資料より筆者作成

## (3) 府省の目標管理型政策評価

府省の政策評価制度のうち、目標管理型政策評価は、各府省の主要な施策を対象として実施される事後評価である。目標管理型評価で作成される評価表には、施策の目的と目標、それら

の達成手段、指標、各手段の目的実現への貢献などの事前の想定を記載する事前分析表と、実施後の測定結果を踏まえて目標達成度を評価する評価書がある（総務省2013、pp.1-3）。

本稿では、事後評価の結果を踏まえた事前分

析表の変更に着目している。事前分析表の項目には達成すべき目標、測定指標、達成手段、その他があり、本稿では測定指標が見直された件数とその理由に着目している。

事前分析表において測定指標が見直された件数を府省ごとに整理すると表2ようになる。指標は必要があれば見直されるものであり、頻繁に見直すことにはデメリットもあるため、府省によって、また年度によって見直しの状況は異なる。また、もともと多数の指標を設定していたか、少数だったかによっても見直しの状況は異なると思われる。よって見直された指標の件数だけでなく、見直しが行われた理由も踏まえる必要がある。

見直された指標とその理由については「各行政機関における政策評価の結果及びこれらの政策への反映状況（個表）」に記載されている。2017年度から2019年度を対象として、測定指標の見直しを行った件数が多い上位8つの行政機関である内閣府、金融庁、消費者庁、総務省、経済産業省、原子力規制委員会、文部科学省、厚

生労働省についてその理由を見ていくと、以下のようにとめられる。

まず、書き方やその詳細さが府省によって異なる。表2のように件数としては整理されていても、見直しの理由については複数の指標についてまとめて記載されている場合や、個別の指標の見直しの理由については省略されている場合などもある。そのため、全ての府省について見直す前の指標、見直した理由、見直した後の指標という一連の情報を整理することはできない。このような状況ではあるが、見直しの理由がわかるものから整理すると、より正確に評価するため、進捗状況をより詳細に把握するため、指標に関連する取組みにおいて変更があったため、指標に関連する取組みが終了したため、目標値を達成したため、有識者からの指摘を踏まえて見直しが必要であると判断したためなどが挙げられる。これらの情報からは、以前の指標のどのような点になぜ問題があり、どのように改善した結果、どのような指標になったのかがわかり、帰納的な情報であると言える。

表2 目標管理型政策評価の事前分析表において見直された測定指標の件数

	2016年度	2017年度	2018年度	2019年度	2020年度	合計	施策の本数（平均）
内閣府	15	5	9	8	0	37	58.2
宮内庁	—	—	—	—	—	—	0.0
公正取引委員会	2	1	1	0	0	4	13.0
国家公安委員会・警察庁	5	2	6	2	0	15	18.0
個人情報保護委員会	1	2	2	5	4	14	5.2
金融庁	20	20	11	9	8	68	12.0
消費者庁	10	9	8	8	10	45	10.2
復興庁	0	0	0	0	0	0	6.0
総務省	8	5	5	6	7	31	18.6
公害等調整委員会	0	1	0	0	1	2	4.0
法務省	0	3	2	0	2	7	28.0
外務省	3	3	4	0	3	13	19.0
財務省	2	0	0	0	0	2	110.6
文部科学省	7	4	2	6	4	23	42.0
厚生労働省	2	4	2	4	10	22	82.6
農林水産省	11	0	0	3	2	16	23.4
経済産業省	15	10	2	9	7	43	27.4
国土交通省	0	0	0	0	0	0	44.0
環境省	0	0	2	2	1	5	10.0
原子力規制委員会	3	6	5	6	6	26	5.8
防衛省	0	0	0	0	0	0	24.2
合計	104	75	61	68	65	373	550.4

単位：件

(注1) —は、評価対象となる施策がなかったことを意味している。

(注2) 指標は1つの施策に複数設定されることが多い。そのため、指標の数と施策の数は1対1で対応するわけではないが、参考情報として施策の平均本数も記載している。

(注3) 施策の本数の平均値は2016年度から2020年度の各府省の基本計画または実施計画からカウントして算出している。

(出所) 各年度の「政策評価等の実施状況及びこれらの結果の政策への反映状況に関する報告」より筆者作成

## 5. 指標の質向上の方法と特徴

指標に改善が必要である場合、その指標を見直すことが直接的な方法の1つだが、適切な指標を設定できるようにするという方法も一般的なものの1つである。4では、指標の見直しに関する現状を整理し、見直しから得られる情報は帰納的な情報であることを確認した。5では、適切な指標を設定できるようにするための取組みとして、外部評価によるチェック、マニュアルやガイドラインの作成、評価研修を通じた学習を取り上げる。その上で、これらの方法から得られる情報には帰納的な情報が不足していることを述べる。

### (1) 外部評価によるチェック

第1は、外部評価によるチェックである。評価制度を導入している地方自治体では47.3%が実施している（総務省 2017、p.3）。地方自治体の外部評価の主要な目的は、自己評価の妥当性や客観性、正確性の確保や、評価制度の適切な運用の確保である。指標の妥当性のチェックは、自己評価の妥当性や客観性、正確性の確保に関連するものである。外部評価では、指標に問題がある場合、代替案が提案される場合もあるが、自己評価の妥当性をチェックするという趣旨からも、不適切な指標であることの指摘や、どうあるべきかということに関する発言が中心であると考えられる。例えば、1つの地方自治体のみを対象とした研究ではあるが、地方自治体の外部評価の傍聴と議事録から、自己評価についてどのような指摘がなされているかを明らかにしたものがある（池田 2019b）。その研究では、調査対象とした期間における全指摘の約40%は指標に関するものであり、それらの指摘の中では、アウトプットやアウトカムを把握するための指標として不適切であり、評価に必要な情報が得られていないという指摘が最も多いとされている（池田 2019b、pp.159-160）<sup>1</sup>。

よって外部評価によるチェックから得られる情報は、どちらかというと帰納的な情報だが、改善すべき、あるいは問題があるなどの判定のような情報に留まる場合や、外部評価者の主観

に基づく指摘の場合もある。また、外部評価者の属性は地方自治体によって異なり、委員の属性によって発言の内容も異なるため、指標に関する指摘が必ずしもなされるわけではない。例えば池田（2019b）が調査対象としている地方自治体の外部評価委員会は委員の半数が学識経験を有する者である。また、公共政策の評価に詳しい委員も含まれていることから、例えば公募や無作為抽出で選ばれた市民が多い委員会や、公共政策の評価に詳しい委員がいない委員会などと比べると、指標に関する指摘がなされやすいと考えられる。

### (2) マニュアルやガイドラインの作成

第2は指標の設定についてのマニュアルやガイドラインの作成である。マニュアルを作成して公開している地方自治体は多くないが、例として京都市や枚方市が挙げられる（京都市 2011；枚方市 2021b）。この2つの地方自治体が作成しているものは、構成も内容も同じであり、詳細は4で述べたとおりである。京都市は2004年度から政策評価を実施しており、適切な指標の設定と見直しのために「客観指標の設定マニュアル」を2011年度に作成した。枚方市は2000年度から事務事業実績測定を実施してきたが、2019年度にロジックモデルや指標、目標のブラッシュアップを図るために評価制度の見直しを実施した（枚方市 2021a、p.2）。この評価制度の見直しにあたって評価表の記載項目を変更したことから「事務事業効果測定指標設定ガイドライン」を2020年度に作成した。これらのマニュアルでは、指標や目標値の必要性と、ロジックモデルを用いた指標の設定方法、指標の妥当性をチェックする際の視点が具体例も用いて説明されている。また、効果測定を実施した結果、設定したロジックモデルのとおり効果が生じていない場合には、ロジックモデルや指標、目標などを見直す必要があることについても説明されている。なお、京都市と枚方市のマニュアルはそれぞれのホームページで公開されており、誰でも閲覧し、利用できる。

地方自治体については、このような各地方自治体が作成しているものではなく、地方自治体

を対象として国が作成しているものもある。その代表的な例としては、KPI（Key Performance Indicator：重要業績評価指標）の設定が重視されている地方創生に関するものが代表例として挙げられる。「地方創生事業実施のためのガイドライン・事例集」が作成されており、その1つとしてKPIを設定する際の視点を説明したものがある。このガイドライン・事例集では、KPIを設定する際に注意すべき点が説明されている。また、地方創生の基本目標ごとに取り組みと指標の例が紹介されており、分量的には事例集の方が多い。なお、このガイドライン・事例集は地方創生の内閣官房・内閣府総合サイトで公開されており、誰でも閲覧し、利用できる。内容としては地方創生に関するものだが、地方創生では幅広い分野の政策が扱われているため、地方創生に関する政策以外でも役立つと考えられる。その他に、個別の政策分野における評価のために作成されているものも挙げられる（例えば環境省水・大気環境局水環境課 2009；厚生労働省保険局・健康保険組合連合会 2017；厚生労働省老健局老人保健健康課保健指導係 2002；国土交通省総合政策局安心生活政策課 2018；国土交通省都市局都市計画課 2014など）。

府省の政策評価制度では、各府省の基本計画と実施計画においてどのような指標を設定すべきかということが規定されている。これらは、マニュアルやガイドラインほど詳細ではないが、どのような条件を満たしているべきなのか、どのような要素が必要なかわかるようになっている。例えば4で取り上げた8の行政機関の実施計画では、検証性の高さ、具体性、客観性などの具体的な要素が挙げられている場合もあるが、特に詳しくは説明されていない場合が多い。また府省の目標管理型評価では「目標管理型の政策評価の実施に関するガイドライン」も重要であり、目標と指標の重要性には言及しているが、目標の達成度合いを測定するものが指標であるという説明に留まっている。

また、指標の設定に類似する情報を得られるものとして、ロジックモデルの作成に関するマニュアルやガイドラインが挙げられる（例えば日本財団 2017；PwCあらた有限責任監査法人

2017など）。ロジックモデルの明確化と指標の設定は関連性が高いため、ロジックモデルに関する資料では指標についても言及されている場合が多い。

マニュアルやガイドラインでは、指標の例も示されているが、どのように設定すればよいか、どのような点に注意すべきか、また望ましい指標とはどのようなものかを説明するものである。よってマニュアルやガイドラインから得られる情報は演繹的なものであると言える。

### (3) 評価研修を通じた学習

第3は評価研修を通じた学習である。評価研修とは、評価の質向上や、評価を行う上で必要な職員の知識の増加・深化や技術の向上を目的として実施される研修である。

評価研修は地方自治体や総務省行政評価局などの行政機関だけでなく、行政職員を対象とした研修を提供している外部の主体が実施している場合もある（池田 2019a, pp.87-89）。内容は講義中心のものや事例紹介、実践的なものなど様々である（池田 2019a, pp.87-89）。業績測定は広く用いられている評価手法だが、全ての評価研修で指標に関する内容が含まれているわけではない。しかし、指標の設定は評価に関する課題や困難な点として挙げられることが多いことを踏まえると、テーマとして取り上げられる頻度は比較的高く、ニーズも一定あると考えられる<sup>2</sup>。

ただし、その時々注目されているテーマに影響される部分もある。例えば、総務省行政評価局が実施している統一研修のテーマは以下のようになっている。中央研修で2015年度から2020年度に実施された20の講義のうち主要なものはEBPMに関するものが8、ロジックモデルに関するものが3、データの利活用に関するものが2である。地方研修で2015年度から2020年度に実施された185の講義のうち主要なものはEBPMに関するものが40、データの利活用に関するものが29、指標に関するものが15、ロジックモデルに関するものが7である<sup>3</sup>。このように、指標に関することが中心的なテーマである研修は少なく、別のテーマの研修の一部として扱われる場合が多く、割り当てられる時間としては長くはないと

考えられる。

評価研修において、指標が取り上げられる場合には事例が示されることも多い。また、実践的なワークが取り入れられることも増加しており、実際に指標を設定してみる、ロジックモデルを作成してみるなどの時間が設けられる場合もある。例えば統一研修の地方研修では、2015年度から2020年度に実施された185の講義のうち32が演習型であり、実践的な要素が取り入れられたものである。

なお、評価研修は実施していないが評価表の作成に関する説明会を毎年度実施している場合はある。しかし、説明会は評価表の作成という実務上の作業の説明である場合が多く、知識やスキルの向上という側面は弱いと考えられる。

評価研修を通じて得られる情報は、どのようにすれば指標を設定できるか、指標はどうあるべきなのかといった理論や考え方、適切な指標のあり方に関する情報が中心である。よって、これらは演繹的な情報であると言える。

## 6. 指標の設定に関する帰納的な情報の不足とその必要性

4では指標の見直しに関する現状を整理し、5では指標の見直し以外の方法で、指標の改善のために実施されているものを3つ取り上げ、それらから得られる情報の特徴を述べた。6では5までの内容を踏まえ、指標の設定に関する帰納的な情報の不足とその必要性について述べる。

4では、改善の必要性が明らかになった指標について、どの程度どのように見直されているのかを整理した。本稿では京都市の政策評価制度と府省の目標管理型評価を取り上げたが、指標が見直される理由については、府省の目標管理型評価と京都市の政策評価制度の間に大きな違いはなかったと言える。主な理由は、評価をより有効に行うため、関連する法律の改訂や計画の改定に対応するためといったものである。ただし、評価をより有効に行うためという理由については、もとの指標では有効に評価できない理由、新たな指標によって有効に評価できるよ

うになることについての説明などもう一段階詳細な情報もあれば指標の改善により役立つと考えられる。指標を見直した理由としては同じであっても、指標が設定される公共政策については実施内容や手法などが多様だからである。

5の(1)では外部評価によるチェックから得られる情報を取り上げ、改善すべき、あるいは問題があるなど判定に近いが帰納的な情報であることを明らかにした。ただし、地方自治体では、外部評価の実施方法が異なり、外部評価者の属性によっても得られる情報は異なると考えられる。5の(2)ではマニュアルとガイドライン、5の(3)では評価研修から得られる情報を取り上げた。これらの方法から得られる情報は、指標はどのようなものであるべきか、どのような手順で設定するのかといった理論的なものや多くのものに共通する一般的なものが中心であり、これらは演繹的な情報であると言える。このように、指標の改善のために実施されている主要な方法から得られる情報は演繹的なものが中心である。2で述べたように、事前に設定した指標を用いて評価するという業績測定の特性上、指標に関する理論を理解した上で適切なものを設定できるようになることの必要性や重要性は高い。しかし実態としては、指標の不適切さが指摘されることや指標設定の難しさが課題として挙げられることが多いという状況は続いている。そのため、既存の方法に加えて帰納的な情報を活用していくことも必要であると考えられる。

このような帰納的な情報の必要性、重要性は指標の設定にのみあてはまることではない。公共政策については事例集が作成されることが多く、これは帰納的な情報の提供方法の1つであると言える。例えば様々な交付金において、要綱や要領に加えて事例集が作成されている場合が多い。前者はどのような政策であるべきかなどの演繹的な情報、後者はどのような取組みが実施されているかなどの帰納的な情報を提供するものであると言える。事例集は望ましい事例の普及やアイデアの提供・相互参照などを目的としており、優れたものや注目すべきものを取り上げる場合が多いが、問題のあるものを取り

上げているものもある<sup>4</sup>。

また、設定した指標を用いて評価をすれば、問題があることはわかる場合もある。例えば想定以上に測定に手間がかかる、あるいは難易度が高い、測定はできるがデータを入手できる時期が評価のスケジュールと著しく合っていない、外部評価で問題を指摘されるなどが挙げられる。このような情報を収集し、それらを蓄積することで適切な指標と不適切な指標、指標を見直すべきタイミング、見直す方法に関する理解が深まり、適切な指標に近づけていくことができると考えられる<sup>5</sup>。

加えて、指標の設定も含めて、評価については継続的な学習・訓練の必要性が指摘されている（稲沢 2021；北大路 2004、pp.316-318；田中 2014、pp.290-294；古川 2008、p.150）。指標の見直しに取り組み、その状況を整理することで、指標の質改善のための既存の取組みから得られる演繹的な情報に帰納的な情報を加えることができる。それによって改善や学習の効果が高まることが期待できる。見直される前の指標と見直しの理由、見直し後の指標という一連の情報がわかることで、適切な指標と不適切な指標について学習できる。また、担当者が変わっても、なぜこの指標が設定されているのか、どのような場合に指標を見直す必要があるのかわかることは役立つと考えられる。行政では異動が多いことから、自ら新しい指標を設定する機会よりも、以前の担当者が設定した指標をそのまま用いるか、見直すかを考える機会も多いと考えられる。そのため、指標の設定や見直しに関する情報が整理されていることは役立つと考えられる。また、地方自治体や国の行政職員は全員が完璧な指標を設定できるようになってから評価に関わるわけではない<sup>6</sup>。そのため、理論を理解した上で指標を設定するという演繹的な方法に加えて、事例から学習していく帰納的な方法も必要ではないかと考えられる。

このように、指標の見直しに関する情報は自己評価の質向上やアカウンタビリティの観点から重要性が高いと言える。本稿で整理してきたように、指標の質改善の方法としては、不適切であることがわかった指標を見直すことと、適

切な指標を設定できるようにすることの2つのアプローチが考えられる。現在の取組みとしては後者を中心ではあるが、指標の見直しから得られる帰納的な情報から学習できるようにすることも指標の質改善に必要であると考えられる。

## 注記

- 1 その他の指摘としては、記入漏れや数値の転記ミスなどの記載ミス、何を意味する指標なのかわからない、目標値が設定されていないなどが挙げられている（池田 2019b、pp.159-160）。
- 2 地方自治体における評価研修に関する研究や公表されている情報は少ない。そのため必ずしも一般化はできないが、研修に関する調査を行った研究では、行政職員からのニーズが高いテーマであるという結果を得ている（池田 2019a、pp.93-94）。
- 3 1つの研修について、重複して複数のテーマにカウントしている場合もある。
- 4 例えば、地方創生交付金の不採択事例について、事例ごとにその要因を分析したものが挙げられる（内閣府地方創生推進室 2015）。
- 5 地方創生に関する取組みの有効性を向上させるために、その内容や指標や目標値、成果について地方自治体が相互に参照・学習することが必要であるという指摘もある（早川・金崎・北山 2021、p.749）。
- 6 産業振興に関する部署の職員のみを対象とした調査ではあるが、業務に関連する分野についてどの程度詳しいかについては、勤務年数や最終学歴が影響を与えているという調査結果もある（小川・津布・家森 2016、pp.11-13）。この他に、異動が多いことも関係していると考えられる。つまり、必ずしも準備が整ってからその業務に取り組んでいるわけではないと考えられる。

## 参考文献

- 池田葉月（2019a）「評価研修におけるワークの難しさと座学の重要性－評価の改善によるよい政策の実現を目指して－」、『福祉社会研究』、(19)：85-98
- 池田葉月（2019b）「評価表のチェックリスト活用による自己評価のレベルアップ」、『日本評価学会第20回全国大会発表要旨集録』、155-160

- 稲沢克祐 (2021) 「わが国自治体における業績マネジメントー行政評価を活用した成果志向の予算編成ー」、『評価クォーター』、(59) : 37-55
- 上野宏・上野真城子 (2008) 「パフォーマンス・メジャーメント」、三好皓一『評価論を学ぶ人のために』、世界思想社
- 海川能理子 (2021) 「EBPMとロジックモデルは難しい」、佐藤徹『エビデンスに基づく自治体政策入門ーロジックモデルの作り方・活かし方』、公職研、25-36
- 小川光・津布久将史・家森信善 (2016) 『地方自治体職員から見た地方創生の現状と課題ー産業振興行政担当者に対する意識調査の概要ー』
- 小野達也 (2020) 「業績測定と評価指標」、『プログラム評価ハンドブックー社会課題解決に向けた評価方法の基礎・応用ー』、晃洋書房、177-187
- 環境省水・大気環境局水環境課 (2009) 『水辺のすこやかさ指標 (みずしるべ)ー水環境健全性指標 (2009年版)ー』
- 京都市 客観指標の設定マニュアル  
(最終閲覧日: 2022年1月12日)  
<https://www.city.kyoto.lg.jp/sogo/cmsfiles/contents/0000035/35589/manual.pdf>
- 京都市 行政評価条例  
(最終閲覧日: 2022年1月12日)  
<https://www.city.kyoto.lg.jp/sogo/page/0000036091.html>
- 京都市 政策評価  
(最終閲覧日: 2022年1月12日)  
<https://www.city.kyoto.lg.jp/sogo/page/0000035589.html>
- 行政機関が行う政策の評価に関する法律  
(最終閲覧日: 2022年1月12日)  
<https://elaws.e-gov.go.jp/document?lawid=413AC0000000086>
- 窪田好男 (2014) 「自治体評価の深化と進化ー業績測定から聞き取りへー」、『評価クォーター』、(28) : 14-28
- 厚生労働省保険局・健康保険組合連合会 (2017) 『データヘルス計画作成の手引き (改訂版)』
- 厚生労働省老健局老人保健健康課保健指導係 (2002) 「保健事業評価マニュアル」  
(最終閲覧日: 2022年1月12日)  
<https://www.mhlw.go.jp/topics/2002/03/tp0307-2.html>
- 国土交通省総合政策局安心生活政策課 (2018) 『観光地のバリアフリー評価マニュアル』  
国土交通省都市局都市計画課 (2014) 『都市構造の評価に関するハンドブック』  
政策評価審議会 (2021) 「政策評価審議会提言」  
総務省 政策評価等の実施状況等の国会報告、政策評価結果の予算要求等への反映状況  
(最終閲覧日: 2022年1月12日)  
[https://www.soumu.go.jp/main\\_sosiki/hyouka/seisaku\\_n/nenji\\_houkoku.html](https://www.soumu.go.jp/main_sosiki/hyouka/seisaku_n/nenji_houkoku.html)
- 総務省 政策評価に関する研修等  
(最終閲覧日: 2022年1月12日)  
[https://www.soumu.go.jp/main\\_sosiki/hyouka/seisaku\\_n/seisaku\\_forum.html](https://www.soumu.go.jp/main_sosiki/hyouka/seisaku_n/seisaku_forum.html)
- 総務省 (2013) 「目標管理型の政策評価の実施に関するガイドライン」  
総務省自治行政局市町村課行政経営支援室 (2011) 「地方公共団体における行政評価の取組状況」  
総務省自治行政局市町村課行政経営支援室 (2014) 「地方公共団体における行政評価の取組状況等に関する調査」  
総務省自治行政局市町村課行政経営支援室 (2017) 「地方公共団体における行政評価の取組状況等に関する調査」
- 田中啓 (2014) 『自治体評価の戦略』、東洋経済新報社  
坪井秀次 (2021) 「VUCA時代に適合する自治体の新たなマネジメントサイクルーPDCAサイクル・OODAループの比較とFSマネジメントサイクルの適合性についてー」、『評価クォーター』、(59) : 21-36
- 内閣府地方創生推進室 (2015) 「地域活性化・地域住民生活等緊急支援交付金 (地方創生先行型) 先駆的事業分 (タイプ I) の不採択事業の要因分析について」  
内閣府地方創生推進事務局 (2020) 「地方創生推進交付金の活用事例に関する調査・分析について」  
日本財団 (2017) 『ロジックモデル作成ガイド』  
ハトリ H. (2004) 『政策評価入門』、東洋経済新報社  
早川有紀・金崎健太郎・北山俊哉 (2021) 「地方創生政策の特徴と課題ー関西2府4県自治体アンケート調査をもとにー」、『法と政治』、72 (2) : 743-766
- 枚方市 (2021a) 『事務事業効果測定指標評価報告書』  
枚方市 (2021b) 『事務事業効果測定指標設定ガイドライン』  
PwCあらた有限責任監査法人 (2017) 『ロジック・モ

## デル作成の手引き』

- 古川俊一（2008）「自治体評価」、三好皓一『評価論を学ぶ人のために』、世界思想社、134-150
- 古川俊一・北大路信郷（2004）『新版 公共部門評価の理論と実際－政府から非営利組織まで－』、日本加除出版
- 三菱UFJリサーチ&コンサルティング自治体経営改革室（2021）『令和2年度自治体経営改革に関する実態調査 報告書』
- 三好皓一（2008）『評価論を学ぶ人のために』、世界思想社
- 山谷清志（監修）（2020）『プログラム評価ハンドブック－社会課題解決に向けた評価方法の基礎・応用－』、晃洋書房
- 米原あき（2020）「データの収集・分析1－量的データ分析－」、『プログラム評価ハンドブック－社会課題解決に向けた評価方法の基礎・応用－』、晃洋書房、147-159
- 龍慶昭・佐々木亮（2010）『「政策評価」の理論と技法 増補改訂版』、多賀出版

（2022.6.10 受理）

## **The need of inductive information to improve the quality of indicators in evaluation by performance measurement**

Hazuki Ikeda

Kyoto Prefectural University  
ikedahazuki@kpu.ac.jp

### **Abstract**

Performance measurement is a relatively simple evaluation method, and evaluation systems that use performance measurement as their central method are widespread in the public sector. Since performance measurement is a method of evaluation by comparing actual values measured using predetermined indicators with target values, the quality of the indicators is important. Therefore, it is necessary to set appropriate indicators and review inappropriate ones. There are various ways to improve the quality of indicators, one of which is to organize and learn from the information on the indicators before the review, the reasons for the review, and the indicators after the review. However, this paper finds that there is little information on the status of implementation of the review of indicators and the reasons for the review. It also reveals that there is a lack of inductive information, such as information on the review of indicators, and describes how this information can help improve indicators.

### **Keywords**

performance measurement, setting and reviewing indicators, improving the quality of self-evaluation

## 【研究ノート】

## ウランバートル市の大気汚染対策の政策評価 ーロジックモデルの観点からー

オンドラハ バトホヤグ

京都大学地球環境学舎

batkhuyag.undrakh.66z@st.kyoto-u.ac.jp

### 要 約

2007年以来、モンゴルの首都ウランバートル市において大気汚染が社会問題となった。その原因の8割はゲル地区の暖房兼キッチン用燃料（石炭）である。2018年、モンゴル政府は、決議第62号「原炭の使用禁止」で、石炭使用を禁止し、改良燃料を提供した。

本研究の目的は、ロジックモデルをウランバートル市の大気汚染対策に適用し、実態を検討して課題を明らかにすることである。決議第62号のニーズ評価とプロセス評価を行った。プロセス評価で、各政策の因果関係を明らかにし、各政策のアウトカムのKPI指標を設定し、効果を検証した。最後にアウトカムの指標を比較した上で、決議第62号について実施前と実施後の大気質と呼吸器疾病率を比較し、本決議の利点と改善点を明確にした。

### キーワード

政策評価、ロジックモデル、大気汚染、ウランバートル

## 1. はじめに

### (1) 評価とロジックモデル

評価とは、政策プログラムの改善に寄与するための手段として、明示的または黙示的な基準と比較しながらプログラムや政策の実施あるいはアウトカムを体系的に査定することである（ワイス 2014：5）。更に、体系的な情報を収集し現在の事実を特定した上で価値判断をすることである（佐々木 2020：3）。プログラム評価の5階層または行政の活動が最終的な成果につながるまでの因果関係を論理的に図式化した重要な評価の手法として、ロジックモデルがある。ロジックモデルとは、プログラムを運用するために必

要なリソース、計画するアクティビティ、達成したい変化や結果の間の関係に関する理解を提示および共有するための体系的かつ視覚的な方法である（Kellogg 2004：1）。

ロジックモデルに関する先行研究は、①個別プログラムへのモデルの適用、②モデルの効果、③モデルの構築方法や利用方法の体系的整理に分類される（佐藤 2017：3）。ロジックモデルはさまざまな社会問題の政策に適用されてきた。

世界で最も大気汚染が深刻視されているのは、バングラデシュの首都ダッカ、パキスタンのラホール、インドの首都ニューデリー、モンゴルの首都ウランバートルである。大気汚染の原因である化石燃料の排気ガスの多くは、COやNO<sub>2</sub>

であり、気候変動と人間の健康に大きく影響を与えることで知られている。世界全体では、2016年には、家庭由来の大気汚染が380万人の死亡の原因となり、世界の死亡率の7.7%を占めており、その影響は深刻である（WHO 2021）。

このようなプログラム評価の発展したツールは、ウランバートル市の政策評価においても有用であると考えられる。

## (2) 先行研究

ウランバートル市の大気汚染に関する調査報告や研究を概観する。世界銀行（The World Bank 2011）は、①国際的な大気汚染削減マネジメント方法、②世界的に成功した大気汚染削減対策、③現在のウランバートル市における大気汚染の被害と大気汚染発生率を評価し、各発生源から排出される大気汚染物質の割合を数値化した。また、大気汚染削減に必要な予算の基準を設定した。大気汚染発生源の調査、発生源の抑制、管理能力向上に関するプロジェクトも行われている（国際協力機構 2017）。それに加えて、即時に実施が必要とされる総合的な対応や、法的環境および規制の枠組みの改善に関する報告書（UNICEF 2016）がある。

学術研究としては、ウランバートル市の2014年から2020年までの大気中のPM濃度と、原炭から改良燃料への移行によるPMへの影響（Gantuya et al. 2020）、屋内と屋外の大気汚染の関係や、ストーブの燃焼効率向上と上質燃料の効果（Guttikunda et al. 2013）、室内外の大気質（武本ほか 2012）に関する調査が挙げられる。さらに、石炭灰の化学組成・放射線・再利用可能性（Davaabal et al. 2016）、ストーブの便益効果分析（Enkhtsolmon et al. 2016）、大気汚染削減に向けたプログラムの開発および評価（伊藤ほか 2017）に関する研究がある。また、当市の大気汚染シミュレーションも行われている（武本ほか 2015）。

政策評価研究は1960年前後にアメリカで開始され、その後他の国々に広がり、日本で政策評価学が制度化されたのは2000年代である（山谷 2012）。一方、モンゴルは1990年まで社会主義国であったため、国の政策を評価する体制が未整備であった。ようやく2015年に、開発政策、計

画に関する法律が可決されたが、政策評価に関する制度も学術研究も遅れていると言わざるを得ない。

当市の大気汚染対策を政策評価の観点から検討した研究は、管見の限り皆無である。特にロジックモデルを扱った研究はなく、事業の妥当性が十分に吟味されていない。また、現状までのウランバートル市の改良燃料による大気汚染の防止政策の成果を体系的かつ論理的に評価した研究はなく、その有効性は未検証のままである。

## (3) 研究目的と研究方法

モンゴル政府の大気汚染対策の中で、最新のものとして挙げられるのが、モンゴル政府決議第62号「原炭の使用禁止」である。これは、ウランバートル市大気汚染の8割を占める原炭を直接規制するもので、モンゴルの環境汚染対策としては極めて重要な位置を占めるといえる。

本研究の目的はロジックモデルをウランバートル市の大気汚染対策に適用し、実態を検討して課題を明らかにすることにある。

本研究の目的を達成するために、①まず二次資料を用いて決議第62号のニーズ評価を行った。②ニーズ評価を基に、本決議を実施するために策定した各政策を明確にした。そして、プロセス評価で、各政策の因果関係を明らかにし、左から右へ複線フローチャート型ロジックモデルを作成した。③各政策のアウトカムのKPI（Key Performance Index）指標を設定した。これを用いて、各政策の直接成果と中間成果の指標を比較し、各政策の成果を明らかにした。④最後に、決議第62号について実施前と実施後の大気質と呼吸器疾病率を比較し、本決議の利点と改善点を明確にした。

## 2. ロジックモデル

### (1) 理論的概観

ロジックモデルは、プログラムの操作や作業の変更に必要な要素間または因果関係を記述し、共有する方法を提供する。また、政策・プログ

表1 ロジックモデルの種別

モデル名		変化の理論 (TCM-Theory of Change Model)	政策理論モデル (PLM-Program Logic Model)	数量による関連性管理	有効性評価
種類		基本 TCM	基本 PLM		
		非線形 TCM	非線形モデル PLM		総合的評価 <sup>2</sup>
目的		より良いモデルに向けて	より良い政策に向けて	管理のための適宜なフィードバック	<sup>1</sup> 策定者に向けて (改善) <sup>2</sup> 利害関係者に向けて (説明責任)
適用の時期	計画時	第1段階	第2段階	第3段階	
	評価時	第3段階	第2段階	第1段階	

(出所) Lisa and Cynthia (2013) を基に筆者作成

ラムのデザイン、計画、実施、評価のいずれの時点においても用いられる(表1)。本論文では評価の観点からロジックモデルを用いる。

ロジックモデルで分析するには、①エビデンスに基づいており、②信用できて尤もらしく、③戦略と結果が含まれ、④現実的であり、⑤知識と仮定を用いる必要がある。本研究では基本PLM (Program Logic Model) により評価をする。この手法の長所は、次の3点にある。①改善に役立つ情報を提供し、情報を迅速に共有できる。②進捗状況を監視し、必要に応じて途中で修正を行うため、プログラムアクティビティ、アウトプット、および短期アウトカムに最も重点を置くことができる。③改善のための提案の重要性を示すのに役立つ (Kellogg 2004)。一方、注意点は、①簡略化された形式であるため、簡潔さと複雑さのバランスをとるのが難しいこと、②適切な境界を確立するのは難しい中で介入全体を理解する能力を保つ必要があること、③定期的に途中のレビュー・更新が必要であること、④創造的思考をしなくなる可能性があるため、柔軟な使用が必要であること、⑤より強力な方法にするためには他のツールと組み合わせる必要があること (Petersen et al. 2013 : 7)。

## (2) 事例研究

ロジックモデルを用いて政策の因果関係を明確化し評価した日本での研究としては、日本の総務省がロジックモデルの下で統計的分析手法

を用いて500の政策を評価した事例研究がある(総務省行政評価局 2019)。また、ロジックモデルを策定した後のアウトカムセッションの評価における指標に関して(茂木 2016)、政策の実効性を向上させるためのロジックモデルを用いた最終アウトカム・業績測定の評価指標設定が研究されている(小野 2018)。

## 3. 大気汚染と対象地域

### (1) ウランバートル市とその特徴

ウランバートル市は標高1,320メートルに位置し、比較的緯度が高く、海岸線から千キロメートル以上も離れている盆地である。平均気温は、春5℃、夏30℃、秋7℃、冬-25℃になるため、世界で最も寒い首都と言われている。夏は短く涼しいが、冬は長く寒い上に乾燥しており、年間平均降水量は242.7mmであり、平均湿度は69%である。さらに、地理的な位置と気候は同市の大気の質に大きな影響を与えている。冬の空気はシベリアの高気圧の強い影響を受けており、地面の冷たい空気層が温かい空気層に覆われ、汚染された大気が盆地内の同市に停滞してしまう (NSO 2020)。

ウランバートル市では、2007年頃から大気汚染が社会問題化している。その背景として、2002年6月27日、モンゴル国民に土地所有を認める土地所有法が成立したことが挙げられる。この法

律により、モンゴル国民は、ウランバートル市に7アールまでの住宅用地を無料で取得することができるようになり、この利用権は2028年5月1日まで有効とされている。この結果、同市の人口集中やゲル地区の拡大が急速に進んだ。

2020年時点で、モンゴルの全人口の46.08%が首都ウランバートルに住んでおり、その96.3%は中心地の6地区に住んでいる（NSO 2020）。同市の人口は2000年から2014年までに2倍に増加し、その大きな原因は地方からの人口移動である。この人口移動により、ゲル地区の占める面積は、市の40%以上に拡大した（Baasantogtokh 2013）。ゲル地区住民の増加に伴い、汚染状況はますます悪化している。2017年の調査では、冬期（10月から4月までの期間）におけるウランバートル市の大気汚染は、北京、ニューデリーなどの工業化された世界の大都市を上回り世界最悪となった（Byambatsereen et al. 2018）。冬季の大気汚染源は、4つに大別される。原因の80%は、ゲル地区のストーブと企業が保有する3,000個の加熱ストーブで消費される、年間100万トンもの石炭にある。10%は、10年以上使用されている238,000台の自動車燃料であり、6%は火力発電所の石炭使用、残りの4%がその他の発生源である（統合法律情報センター 2017）。

この大気汚染によって、深刻な健康被害が報告されている。市民の呼吸器疾患が年々増加し、全ての疾病の23.4%～24.4%を占めている（Oyunjargal et al. 2018）。2015年冬季のPM10の日平均値は1.009 mg/m<sup>3</sup>であった。2012年冬季のピーク時の3.000 mg/m<sup>3</sup>からは約3分の1に減少したが、MNS4585：2007<sup>1</sup>の基準値0.100 mg/m<sup>3</sup>の10倍以上であった（数理計画2013、2017）。

大気汚染が同市の社会問題となって以来、モンゴル政府は状況の改善に取り組んできた。なかでも、2018年政府決議である第62号「原炭の使用禁止」は、今までの環境政策の中でも最も効果が高かったと評価されている。

表2 大気汚染による死亡者と死亡率

病気	死亡者	大気汚染が原因である病気の死亡率	
肺炎（0－4歳子供）	435	肺病	34%
心臓病	2151		
循環器系の疾病	1493	肺および心臓病	24%
慢性閉塞性肺疾患	453		
白気管、気管支炎、肺癌	233	その他（全死亡数）	8.2%

（出所）UNICEF（2016）、武本ほか（2012）

## （2）大気汚染防止政策の経緯と現状

ウランバートル市での大気汚染発生に対して、2007年からモンゴル政府が取り組んできたゲル地区に関する政策を整理してみると、次のように要約できる。

2006年、市役所の環境保護局に大気品質部門が設立されたが、2009年に解散し、首都大気質庁が設立された（数理計画2017）。2010年、大気汚染出費法が制定された。この時期は、主に大気質測定や人材育成に注目が集まっていた。

2012年、大気法が制定された。また、市長の命令でバヤンゴル区での調査を行い、同区の原炭を禁止した。2014年、国家大気汚染低減委員会の決議第5号「改善された燃料に関して取られるべき対策」を行った。2014年、国家大気汚染低減委員会が設立された。また、改良燃料の原料になる石炭の各品質が調査され、現存の改良燃料プラントが建設された。この時期はストーブと石炭の調査が注目されていた。

2016年に、衛生法が制定され、各種廃棄物の燃焼が禁止され、2017年には、違反法が制定された。2016年、暖房に原炭を使用する人々に対し電気ヒーターの使用を奨励するために、ゲル地区世帯の夜間電気の価格を引き下げ、電気割引が導入された。政府の決議第98号「大気、環境汚染削減ナショナルプログラム」が制定された。2018年、燃料の要求、需給バランスを調査し調整が行われた。

2016年8月、首都大気庁が解散となり、代わって首都大気汚染削減庁が設立された。また、2017年には、国家大気汚染低減委員会が国家環境汚染低減委員会へと拡大された。

2017年に政府決議第98号「大気、環境汚染削

減ナショナルプログラム」が施行された。2017年から2025年までの実施が予定されている。2019年3月、首都大気汚染消滅庁が解散となり、代わりに首都大気汚染反対庁が設立された（首都大気汚染反対庁 2021）。

2018年2月、原炭の使用を禁止するモンゴル政府決議第62号が制定され2019年5月から実施された。対象は、電気と採熱用に原炭の使用が認可されている組織・企業以外であり、この対象には市民も含まれる。この政府決議の特徴は、改良燃料の価格が原炭の価格を超えないようにし、燃料の需給バランスを確保することである。また、その実施を国家環境汚染低減委員会が担当した。本委員会は、委員長である環境大臣の他、12名の副大臣や市役所・政府機関・NGOの代表など27名から構成されている。2018年、国家大気汚染低減委員会の決議第18/01号により、改良燃料プラントの能力を研究するためのワーキンググループが設立された。燃料分析の結果によると、エネルギー資源株式会社の濃縮石炭の品質は比較的安定していたので、その石炭を利用することにした。また、この会社からの240万トンの石炭とその運搬が無料化された。2018年4月、改良燃料生産が開始された。

国際機関などによるプロジェクトとしては、2008年、日本の新エネルギー技術開発機構（NEDO）、欧州復興開発銀行（EBRD）、大気清浄基金（CAF）が、モンゴルにおける石炭の品質や改良燃料導入について調査した。2012年から2017年まで、知事によって、ウランバートルの大気汚染を削減するプログラムの一環として、ウランバートル大気清浄プロジェクトが世界銀行からのソフトローンにより実施された。JICAウランバートル市大気汚染対策能力強化プロジェクトも開始され、フェーズⅠ（2010年～2013年）、フェーズⅡ（2013年～2017年）、フェーズⅢ（2018年～4、5年）が行われてきた（数理計画2017）。

市民団体としては、2016年12月に「煙に反対する父母の会」というオンライングループが設立され、また11の練炭生産者が「モンゴル練炭協会」を設立した。

## 4. 分析と評価

本研究では、まず、ロジックモデルを用いて政策のニーズ評価とプロセス評価をする。その上で、評価指標を設定し、各政策の有効性と総合対策の限界を明らかにする。

ロジックモデルを作成する際には、Lisa and Cyntia 2013を参考にする。同本では、ロジックモデルに基づいた作成方法を示しており、特に評価のためのロジックモデルを政策の管理に結びつけることの有効性を示している。

### (1) ニーズ評価

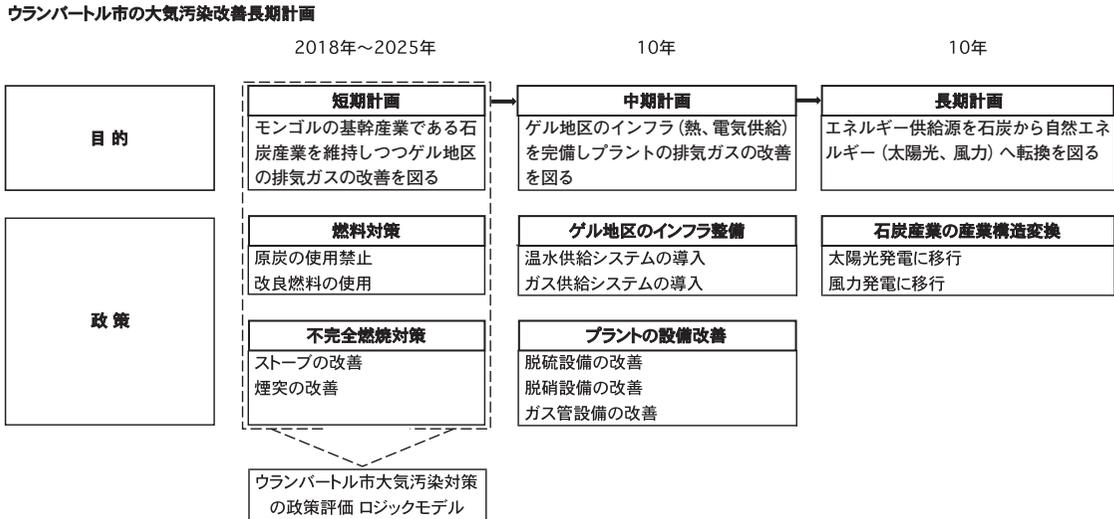
政府の決議第98号「大気、環境汚染削減ナショナルプログラム」（2017）とJICA「ウランバートル市大気汚染対策能力強化プロジェクト」フェーズⅡの報告書（2017）を基にニーズ評価を行った。図1で示しているように、2018年から2025年まで、短期計画としてモンゴルの基幹産業である石炭産業を維持しつつ、ゲル地区の排気ガスの改善を図ることを目的としている。2025年までに、2016年の大気汚染の基本レベルと比較して80%削減することを目指している。そのための対策として、石炭の使用を禁止し、改良燃料の使用を開始することになった。不完全燃焼対策としてストーブと煙突を改善した。この短期計画を実施するため、モンゴル政府決議第62号が制定された。本研究の対象となる政策であり、本決議が実施される地域を図2で示した。

中期計画として、ゲル地区のインフラである熱、電気供給を完備しプラントの排気ガスの改善を図る。そのため、ゲル地区のインフラ整備とプラントの設備を整える必要がある。長期計画では、エネルギー供給源を石炭から太陽光、風力とした自然エネルギーへの転換を図っている。

### (2) プロセス評価

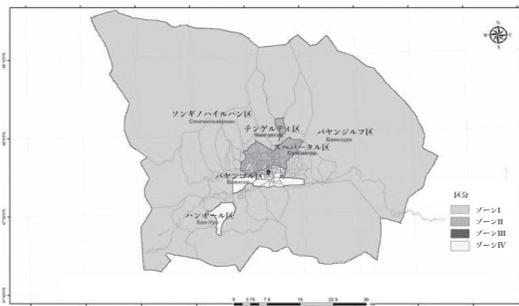
プロセス評価では、政府決議第62号とそれに関連する各政策と結果の因果関係、評価指標を設定した。改良燃料を導入するまでの段階（図3）・導入後の段階（図4）に分けて実施された関連政策の抽出とその政策の因果関係を、エビデ

図1 決議第62号のニーズ評価のロジックモデル



(出所) 筆者作成

図2 ウランバートル市の空気質改善ゾーン



(注) 2020-2021年の空気質改善ゾーンで、Iは改良燃料を使うゾーン、IIとIIIは電気およびその他の電源を取り組むゾーン、IVは全種類の固形燃料を禁止するゾーン

(出所) 首都大気汚染反対庁(2020)より筆者邦訳

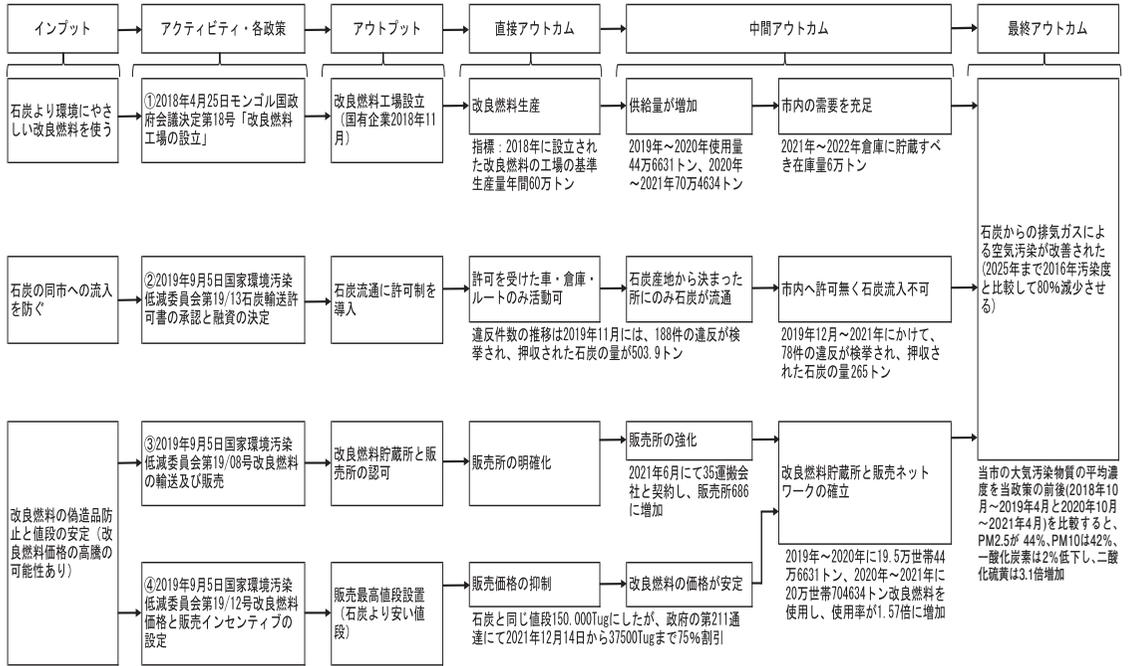
ンスを基にロジックモデルのPLMに適用した。改良燃料を導入するまでの段階と導入後の段階のインプットには、解決する対象となる問題を取り上げ、アクティビティには国家環境汚染低減委員会の政策を取り入れた。アウトプットにはその決定の直後の結果を示した。

アウトカムを直接・中間・最終に分けて各政策のアウトカムをまとめた。なお、政策評価のための評価指標はアウトカムの指標に絞った。改良燃料を導入するまでの段階でのアクティビティには、①改良燃料工場の設立、②石炭輸送

許可書の承認と融資の決定、③改良燃料の輸送および販売、④改良燃料価格と販売インセンティブがある。

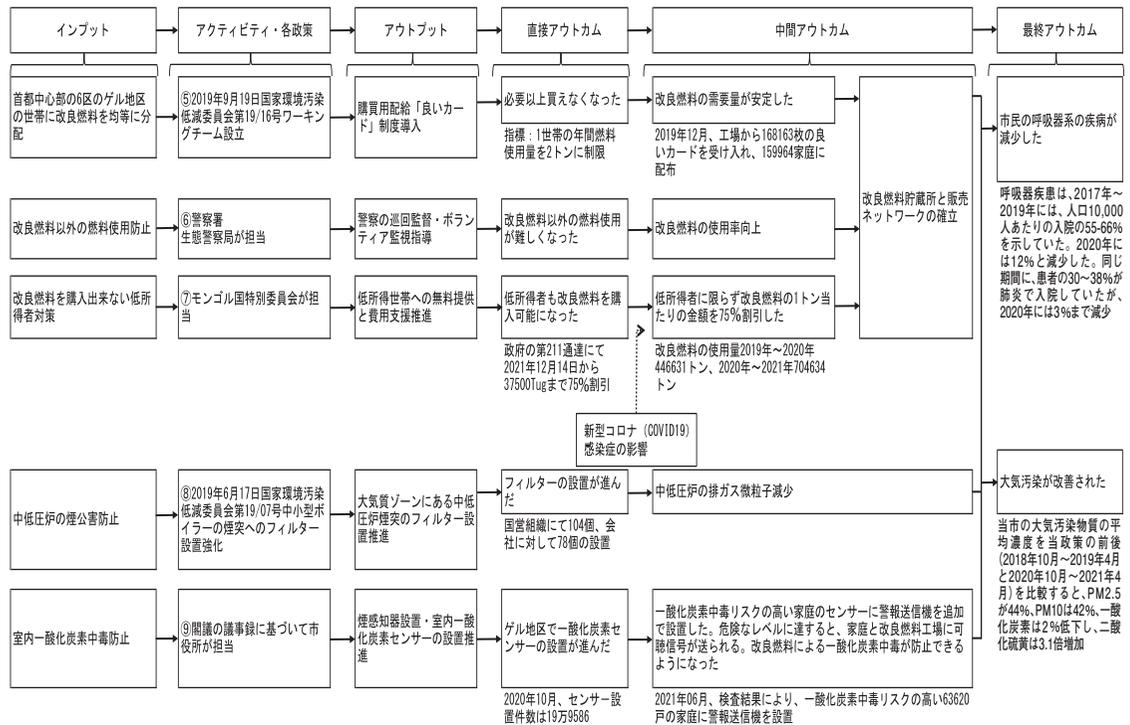
導入後の段階には、計画以外の問題に対して取った取り組みが含まれる。具体的には、以下の5点が挙げられる。⑤改良燃料を均等に分配するため、ワーキングチームを設立し国民による必要以上の買い取りを防いだ。⑥改良燃料以外の燃料の使用を防止するため警察署が巡回監視し、他にボランティアの監視者も参加の上、他の燃料使用を止めた。⑦モンゴル国特別委員会が担当し、改良燃料を購入出来ない低所得者向けに低所得世帯への無料提供と費用支援を推進した。⑧中低圧炉の煙公害を防止し、中小型ボイラーの煙突にフィルターを設置した。⑨室内一酸化炭素中毒防止のため、閣議の議事録に基づいて市役所が担当して煙感知器設置・室内一酸化炭素センサーを設置した。①～④の政策が行われた結果、最終アウトカムとして石炭からの排気ガスによる空気汚染が改善された。⑤～⑨の政策が行われた結果、改良燃料貯蔵所と販売ネットワークが確立されたので、最終アウトカムとして大気汚染が改善され、市民の呼吸器疾病が減少した。

図3 改良燃料の導入までのロジックモデル



(出所) 筆者作成

図4 改良燃料の導入後のロジックモデル

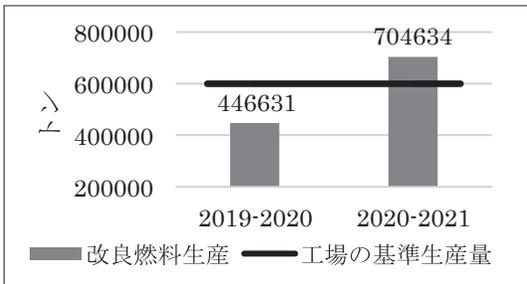


(出所) 筆者作成

(3) 各政策の成果と評価

本節では、図3と図4で示しているアクティビティの①～⑨の各政策の成果を分析・評価する。①の政策は、特定された市民が必要とする改良燃料を確保するためのものである。2018年に設立された改良燃料の工場の基準生産量は、年間60万トンである（図5）。2020年1月には、第2工場の設立が決議され、11月から運営が開始されている。2021年から2022年については、貯蔵すべき倉庫量が6万トンと定められた。工場が増設されたこと、在庫計画があること、1世帯の使用量を制限したこと（政策⑤参照）、ゲル地区の約20万世帯に提供することができるようになったことから、この政策はほぼ目標を達成したと考えられる。

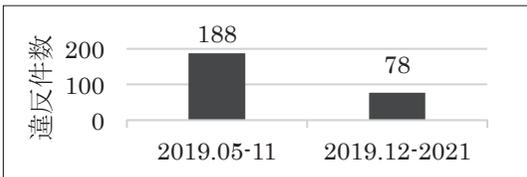
図5 改良燃料生産量



(出所) 首都大気汚染反対庁のデータを基に筆者作成

②に関しては、石炭が同市に流入することを防ぐために石炭流通に許可制を導入したが、違反が発生している。この政策に対する違反件数は、図6で示しているように半減したが、なおも80件近くある（図6）。また、押収した石炭量はゲル地区年間使用量の0.04%程度であるものの、検挙数が全違反件数の氷山の一角だとすると、今後も監視を続ける必要がある。

図6 石炭輸送許可書の承認に対する違反件数

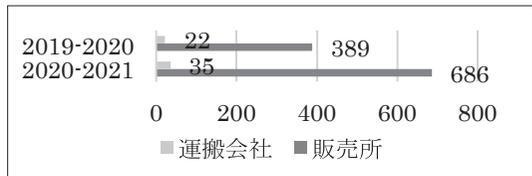


(出所) 首都大気汚染反対庁のデータを基に筆者作成

③の政策では、偽造品防止と販売所を確保するために、改良燃料を運搬する会社と販売所を認可制（図7）にした。

改良燃料の輸送と販売を許可制にすることにより、改良燃料の製造・運送・販売のネットワークが明確になった。

図7 改良燃料の輸搬社数および販売所数



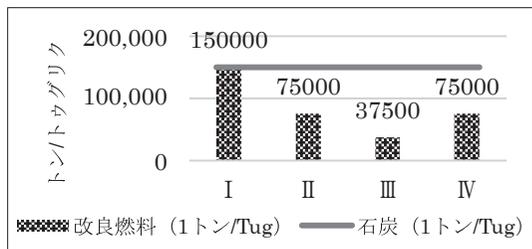
(出所) 首都大気汚染反対庁のデータを基に筆者作成

④については、政府が改良燃料の値段を規定した。工場から販売所までの1トンあたりの運送費を含め石炭と同じ値段にしたが、政府の第211通達によって2021年12月から75%の割引を行った（図8）。その割引が⑦の低所得者向けの費用支援推進にも影響を与えた。

⑤では、改良燃料の不足を避けるため、1世帯あたりの年間燃料使用量を2トンに制限し、それ以上の使用量を必要とする世帯は、個別事情を申告し、それが認められた場合に使用量を増加できるようにした。市民向けの配給制である「良いカード」を導入し、2019年12月の時点では、159,964枚を市役所から市民に配布した。

改良燃料以外のものを使用させないため、⑥に取り組み、首都専門検査機関の専門家111人、郡の職員1,527人、パートタイム雇用の4,227人がパトロールした。

図8 改良燃料価格



(注) I は2019年～2020年12期間、II は2020.12.03時点、III は2020.12.14時点、IV は2021.10.01からの期間を表す

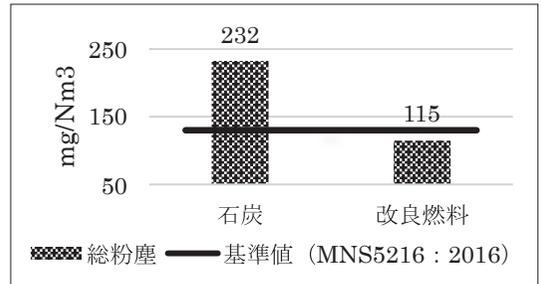
(出所) 首都大気汚染反対庁のデータを基に筆者作成

⑦に関しては、石炭を改良燃料に置き換えるために政府主導で価格を操作してきたが、未だ価格が安定したとは言いがたい。今後は、市場原理も考慮し、価格を安定させる必要がある。

⑧では、国営組織で104個、会社では78個もの中小型ボイラーの煙突へのフィルターを設置したが、空气中微粒子の数値を登録したデータは存在しない。

⑨を評価するため、石炭と改良燃料の排気ガスを比較すると、改良燃料は石炭に比べ、一酸化炭素の排気が10%（図9）、二酸化硫黄の排気が22%増加した（図10）が、質は基準値内であった。逆に基準値を78%も超えていた総粉塵は90%まで減少していた（図11）。このように、国家環境汚染低減委員会は改良燃料のCO、SO<sub>2</sub>が増えることを調査の結果で知ったにもかかわらず、室内一酸化炭素中毒の予防を行わなかったがゆえに、2020年から2021年の冬期に、一酸化

図11 石炭と改良燃料の比較-総粉塵



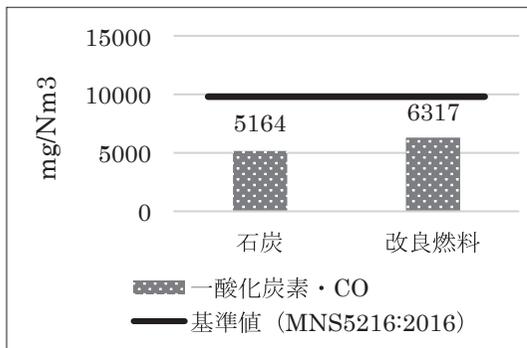
（出所）首都大気汚染反対庁のデータを基に筆者作成

炭素中毒で296世帯859名が病院に搬送され、36名が入院し、2名が死亡した。そこで、各家庭に煙感知器設置・室内一酸化炭素センサーが設置された。当政府決議が実施されている地域内の199,586全世帯に同センサーを設置した上で、室内一酸化炭素中毒に対してリスクの高い63,620世帯を公表した。

ゲル地区ではほぼ全世帯にセンサーを設置し、リスクの高い世帯には通報設備も導入しており、一酸化炭素中毒の防止対策は完成したと考えられる。

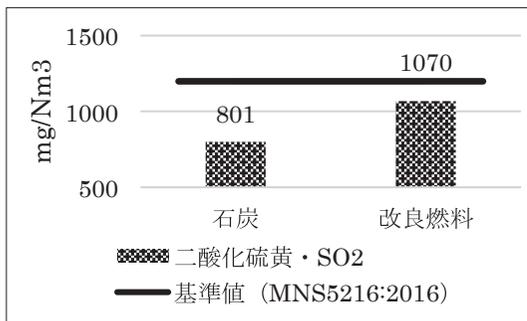
以上の政府対策には一貫性があり、アウトカムを政策実施の前後で比較すると、改善していることが分かる。また、政策を促進する国家環境汚染低減委員会に各省、市役所、NGOをはじめ、合計で27の主体を参加させたことは、広く各立場から政策が看視されることにつながり、効果的であったと考えられる。しかしながら、②～⑨までの各政策の策定時に目的値を定めていないことは問題であった。

図9 石炭と改良燃料の比較-一酸化炭素



（出所）首都大気汚染反対庁のデータを基に筆者作成

図10 石炭と改良燃料の比較-二酸化硫黄



（出所）首都大気汚染反対庁のデータを基に筆者作成

#### (4) 総合政策評価

モンゴル政府の決議第62号の実施に当たって、ロジックモデルで重要と見なされる政策を明らかにし、各政策の因果関係を検証した。主な政策が9つあり、それらの実施期間を改良燃料を導入する前後の2つの段階に分け、それぞれの最終アウトカムを明確化した。また、政策の実施には幅広く、国の機関、自治体、民間が参加しており、各界の意向を十分に反映していた。

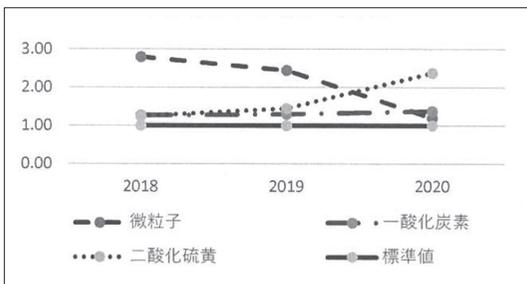
気候・水文学・環境情報研究所によって編集されたデータベースに基づいて、2018年から2020

年の冬期までの大気汚染の程度の推移を、PM10とPM2.5から構成する微粒子、一酸化炭素、二酸化硫黄に着目して分析した。微粒子は当初基準値の3倍近くあったが、基準値近くまで激減している。一方、一酸化炭素は微増したが、1.4倍以下となった。二酸化硫黄は基準値の1.2倍程度であったが、2.5倍近くに激増した(図12)。

冬期の平均気温は2018年には-8.7℃、2019年は-8.1℃、2020年は-10.24℃であった。当市に移動してきた人数は2018年に6,568人、2019年に12,373人、2020年には39,515人である。当市から地方に移住した人数は2018年に6,328人、2019年に6,595人であった(NSO 2020)。

また、微粒子、一酸化炭素、二酸化硫黄のうち2つの変化の相関関係を回帰分析すると重相関Rが0.99～0.988を示しており、大気汚染の変化が同じ要因(石炭から改良燃料への変換)により起こった結果であると推測できる。要約すると、総粉塵量が基準値近くまで削減できたのは成功であり、一酸化炭素も微増に収まっているが、他方で二酸化硫黄が基準値の1.5倍に増加しており、悪化している。二酸化硫黄も健康を害するため、至急改善策を実施する必要がある。

図12 大気汚染度の変化

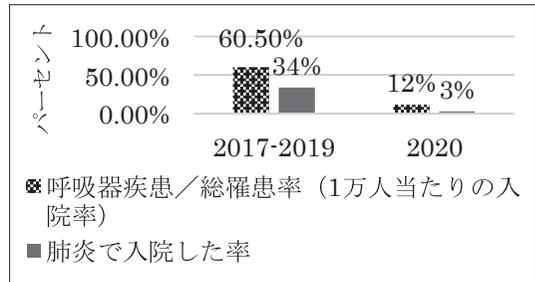


(出所) 首都大気汚染反対庁のデータを基に筆者作成

呼吸器疾患は、2017年から2019年には人口10,000人あたりの入院率の55～66%を占めていたが、2020年には12%と減少した。同じ期間に、患者の30～38%が肺炎で入院していたが、2020年には3%まで減少している(図13)。

言い換えると、大気汚染では二酸化硫黄が基準値の2倍以上に増加しており、その健康被害が懸念されるが、実際には呼吸器疾患の入院率は

図13 呼吸器疾病



(出所) 首都大気汚染反対庁のデータを基に筆者作成

劇的に減少しており、政策の成果は顕著に表れていると考えられる。今後は、二酸化硫黄の市民の健康に対する影響を注意深く見ていく必要がある。

### 5. おわりに

社会問題の原因が明らかであるが、限られた資源しか用いることができない場合、当該問題に対する政策を継続的に改善していくためには、できる限り重複を減らすことが望ましい。そのため、各政策、プロジェクトのアウトカムの策定時に測定可能な数値で目標を明確に表した上で、アクティビティの時点では、数値目標に基づく政策を準備する必要がある。ロジックモデルで各政策の因果関係を検証した結果、モンゴルの大気汚染政策においては、政策の効果は指標で表すことができて、目標値を決めてそれに対する達成率によって表すことは行われていなかったことが明らかとなった。政策の策定時に目的を明らかにし、評価可能な数値で表すことが大きな課題である。

当初貧困対策として、モンゴル政府は、低所得世帯への改良燃料の費用支援を推進しようとしたが、新型コロナウイルス感染拡大の影響で、改良燃料の1トンあたりの金額を75%割引した。それにより、使用量が増加したが、燃料供給会社が割引負担を行っているのが現状であり、取り組みの持続可能性が乏しい。今後の推移が注目される。また、改良燃料から生じる二酸化硫黄の削減方法を探求し、実施する必要性がある。

モンゴルの大気汚染対策の政策評価が少ないなかで、本研究は、ロジックモデルを用いて、初めてニーズ評価とプロセス評価を行った。今後は、費用便益分析などの別の評価方法も活用してゆくことが求められよう。

## 謝辞

本論文の執筆にあたり、貴重なご指導を下さった京都大学大学院地球環境学堂教授の宇佐美誠先生、原稿にコメントして下さった杉谷和哉先生、嶋崎保義様、支援して下さった日本文部科学省（MEXT）、各資料を提供して下さった首都大気汚染反対庁の大気質・監視・管理部長 G.Davaajargal とご担当者の方に感謝したい。

## 注記

1 MNS4585：2007 モンゴルの人々が健康で安全な環境で生活し、働き、勉強するための条件を作り出す、生態系のバランスを確保するために、屋内と屋外の空気中の化学的および物理的汚染物質の許容レベルを決定することが目的である基準。

## 参考文献

伊藤薫・竹内佐智恵・草川好子・田村麻紀・出田慎二・武本行正・高橋正明（2017）「四日市公害の経験を活かしたウランバートル市の大気汚染削減に向けたプログラムの開発および評価」、『四日市看護医療大学紀要』、10（1）：64-72

小野達也（2018）「ロジックモデルを用いた評価指標の設定：業績測定の実効性の向上に向けて」、『評価クォーターリー』、46（7）：46-59

国際協力機構（2017）「モンゴル国、ウランバートル市大気汚染対策能力強化プロジェクトフェーズ3詳細計画策定調査報告書」、<https://openjicareport.jica.go.jp/pdf/12320776.pdf>（2020年03月22日閲覧）

佐々木亮（2020）『評価論理：評価学の基礎』、eBook版、多賀出版

佐藤徹（2017）「わが国自治体におけるロジックモデルの普及実態の分析」、『地域政策研究』、20（1）：1-14

首都大気汚染反対庁（2021）「首都大気汚染反対庁組織紹介」、首都大気汚染反対庁ホームページ  
<https://aprd.ub.gov.mn/taniltuulga>（2021年10月05日閲覧）

数理計画（2013）「モンゴル国ウランバートル市大気汚染対策能力強化プロジェクトフェーズ1」、JICAウェブサイト12088191\_01.pdf（[jica.go.jp](http://jica.go.jp)）（2020年01月22日閲覧）

数理計画（2017）「モンゴル国ウランバートル市大気汚染対策能力強化プロジェクトフェーズ2」、JICAウェブサイト12289187\_05.pdf（[jica.go.jp](http://jica.go.jp)）（2020年01月31日閲覧）

総務省行政評価局（2019）「政策効果の把握・分析手法の実証的共同研究 - 報告書総論ver1.0 -」、  
[https://www.soumu.go.jp/main\\_content/000619709.pdf](https://www.soumu.go.jp/main_content/000619709.pdf)

武本行正・粟屋かよ子・高橋正昭・宮崎真・石川守（2012）「モンゴル・ウランバートル市のPM等大気汚染について」、『四日市大学環境情報論集』、16（1）：21-29

武本行正・高橋正昭・粟屋かよ子（2015）「国内外の環境問題の取り組みモンゴル・ウランバートル市の環境問題と大気汚染の拡散シミュレーション」、『人間と環境』、41（1）：65-75

統合法律情報センター（2017）「大気、環境汚染削減ナショナルプログラム」、<https://legalinfo.mn/mn/detail?lawId=207779&showType=1>（2019年03月10日閲覧）

茂木康俊（2016）「ロジックモデルを用いた政策評価における指標設定方法—自治体レンタサイクル事業を事例として—」、『評価クォーターリー』、36（1）：19-35

山谷清志（2012）『政策評価』、ミネルヴァ書房

ワイス（前川美湖ら監訳）（2014）『入門評価学：政策・プログラム研究の方法』、日本評論社

Baasantogtokh, O.(2013). Ulaanbaatar city: Migration and Issue. *Proceedings of the Mongolian Academy of Sciences*, 53（2）, 51-60.

Byambatseren, Ch., Michidmaa, N., Sonomdagva, Ch., and Yutaka, M.(2018).The Some Results of Study on Outdoor and Indoor Ambient PM2.5. *Proceedings of the Mongolian Academy of Sciences*, 58（1）, 40-54.

Davaabal, B., Battsetseg, B., Zolzaya, Ts., and Temujin, J.(2016).Properties of the ashes from Ger district of

- Ulaanbaatar city and preliminary assessment of their applicability. *Proceedings of the Mongolian Academy of Sciences*, 217, 80-90.
- Enkhtsolmon, O., Matsumoto, T., and Tseveen, E. (2016). Cost benefit analysis of air pollution abatement options in the ger area, Ulaanbaatar, and health benefits using contingent valuation. *International Journal of Environmental Science and Development*, 7(5), 330-334.
- Gantuya, G., Tseren-Ochir, S., and Batbayar, J. (2020). Recent improvement in particular matter (PM) pollution in Ulaanbaatar, Mongolia. *Aerosol and Air Quality Research*, 20, 2280-2288.
- Guttinkunda, K. S., Sereeter, L., Baldorj, B., and Batdorj, D. (2013). Particulate pollution in Ulaanbaatar, Mongolia. *Air quality, atmosphere & health*, 6, 589-601.
- Lisa, W. K., and Cynthia, C. P. (2013). *The Logic Model Guidebook*. Thousand Oaks, Calif: Sage Publications.
- NSO. (2020). National Statistics office of Mongolia. <http://www.en.nso.mn/>. (2020年3月15日閲覧).
- Oyunjargal, D., and Byambatseren, Ch. (2018). The Effects of Environmental Issues on Residential House Price -In the Case of Ulaanbaatar City-. *Proceedings of the Mongolian Academy of Sciences*, 58 (2), 117-130.
- Petersen, D., and Taylor, E. F., and Peikes, D. (2013). *The Logic Model: The foundation to implement, study, and refine patient-centered medical home models*. Rockville, MD: Agency for Healthcare Research and Quality, AHRQ Publication.
- The World Bank. (2011). *Air quality analysis of Ulaanbaatar - Improving air quality to reduce health impacts*. Washington, DC: The World Bank.
- UNICEF (The United Nations International Children's Emergency Fund). (2016). *Understanding and addressing the impact of air pollution on children's health in Mongolia*. <https://www.unicef.org/mongolia/reports/understanding-and-addressing-impact-air-pollution-childrens-health-mongolia>. (2019年11月15日閲覧).
- WHO (The World Health Organization). (2021). *Household Air Pollution Data*. <https://www.who.int/data/gho/data/themes/air-pollution/household-air-pollution>. (2021年4月5日閲覧).
- W. K. Kellogg Foundation. (2004). *Logic model development guide*. Michigan: W. K. Kellogg Foundation.

(2022.6.6 受理)

## **An Evaluation of Air Pollution Control in Ulaanbaatar**

Batkhuuyag Undrakh

Graduate School of Global Environmental Studies, Kyoto University  
batkhuyag.undrakh.66z@st.kyoto-u.ac.jp

### **Abstract**

In Ulaanbaatar, around 80% of air pollution is caused by the use of coal for heating and cooking. In response to this problem, in 2018, the Mongolian government banned the use of coal and allocated substantial resources to the development of refined coal briquettes, as stipulated in Resolution No. 62, “Prohibition of The Use of Raw Coal”.

This study examines the implications of the air pollution control by applying the logic model theory. The application of the theory takes four steps. First, I conducted a need assessment of Resolution No. 62. Second, I implemented a process assessment of the resolution clarifying the causal relationship between each activity and the KPI (Key Performance Index) to examine the outcome of each activity. Third, I compared the KPI between different outcomes. Fourth, I reached the air quality indexes before and after Resolution No. 62 came into effect. The last section is the conclusion and policy recommendations.

### **Keywords**

Policy evaluation, logic model, air pollution, Ulaanbaatar



## 【研究ノート】

## 大学の地域貢献プログラムに質改善を目的とした 評価を行うことの意義

岸本 由梨枝

新見公立大学  
yk6710@niimi-u.ac.jp

### 要 約

近年、大学の地域貢献は、文部科学省による様々な支援策が講じられ全国的に広がっている。本稿は、大学の教育と地域貢献を結びつけ、授業の一環として行われる地域貢献プログラムに焦点を当てている。このような地域貢献プログラムの特徴として、教育と公共政策の側面を持っていることを確認し、評価の現状では、公共政策の視点からプログラムの質を高めるための質改善を目的とした評価は、学生の学習成果に関する研究に比べて数少なく、検証が不十分であることを課題として提示する。また、質改善を図る評価を行うことで、具体的にどのような効果が生じるのか、質改善を目的とした評価を行うことの意義について述べたい。

### キーワード

地域貢献プログラム、公共政策、質改善、プログラム評価、ロジックモデル

### 1. はじめに

現在、大学の地域貢献は全国的に盛んに行われている。18歳人口の減少により、大学は全入時代から淘汰の時代に突入し、地域貢献の取組は大学の生き残りをかけた改革として注目されている。また、政府は多様化し複雑化した社会課題を解決する担い手として大学の役割を期待し、高等教育政策では大学と地域の関わりを強化する様々な支援策が講じられている。

こうした大学の地域貢献は、独立した役割として捉えるのではなく、教育や研究と結びつけて相乗効果を図ることが重要とされており（野澤 2016、p.7）、筆者は特に授業の一環で行われる地域貢献プログラムに関心を持ち、その評価に

関する研究に取り組んでいる。

このような地域貢献プログラムの評価の研究に着目すると、学生が体験を通じて、何を学び身に付けたのか、学生の学習成果に焦点を置いた報告や研究が多い。一方で、地域貢献プログラムを実施したことによる地域への効果や課題に焦点を当てた研究は数少ない（荻野・中川 2021、p.131）。

そこで本稿では、大学の地域貢献が注目された背景とその実態を整理し、地域貢献プログラムは公共政策であり、教育と公共政策の側面を持つことを示す。また、このような地域貢献プログラムの評価の現状を整理し、地域への効果や課題を検証し、質を高めるといった質改善を図る評価の研究は数少なく、検証が不十分であ

ることを課題として提示する。最後に、地域貢献プログラムの質改善を図る評価を行うことで、どのような効果が生じるのか、評価を行うことの意義について述べたい。

本稿で使用する「プログラム」の定義については、山谷（2020）を参考に「社会的目的を達成するために何らかの活動や制度設計が組み立てられたもの」とする。そのため、地域貢献プログラムを「プログラム」と称する。

## 2. 大学の地域貢献の背景と実態

本章では、大学の地域貢献はどのような背景から注目され、様々な支援策が講じられるようになったのか、また、支援策のねらいは何か、授業の一環として行われる地域貢献プログラムはどの程度取り組まれているのか、その背景と実態を整理する。

### (1) 大学の地域貢献の発展と文部科学省による支援策

大学は非営利部門に位置しており、政府部門と協働することや政府部門だけでは対応しきれない社会課題を解決する担い手の一つとして、今後の役割が期待されている。2006年の教育基本法の改正、2007年の学校教育法の改正により、大学は教育と研究の成果を広く社会に提供することで、社会の発展に寄与するとの条項が追加された。これにより、大学の社会貢献は教育と研究に並ぶ第三の使命として位置づけられ、特に文部科学省による大学の地域貢献に対する支援が強化されている。例えば、2008年度から実施された「戦略的大学連携支援事業」（複数大学の連携により大学教育の充実や、地域で活躍する人材の養成を図る取組への支援）や2012年度から2014年度まで実施された「産業界のニーズに対応した教育改善・充実体制整備事業」（地域の産業界と一体となった人材育成や、産業界のニーズを踏まえた取組を行う大学グループに対する支援）等が挙げられる。また、2013年度の「地（知）の拠点整備事業（以下、「COC事業」）」（自治体、地域の企業等との連携のもと、地域を

担う人材を育成するための教育改革や、卒業生の地元定着促進に向けた取組を行う大学に対し、最長5年間にわたる支援）の開始に続き、これを発展させた「地（知）の拠点大学における地方創生推進事業（以下、「COC+事業」）」が2015年度から開始された（野上 2019、p.29）。

例えば、2013年から実施されたCOC事業は、大学全体で地域貢献を進める体制整備を文部科学省が支援した初めての事業である。このような支援策を講じた文部科学省のねらいは、大学を地域の中核的存在として位置づけ、地域に密着した研究や教育を行うことで、地域再生や地域活性化を担う中心的な役割を担わせることである。大学全体で地域貢献の取組を行うことが期待されていた。

### (2) 大学の地域貢献プログラムの実態

前節で述べたように、大学の地域貢献の取組は、大学に対する社会的要請や文部科学省による支援策等の後押しにより広がってきている。では、本稿で取り上げる授業の一環として行われる地域貢献プログラムは、実際にどの程度行われているのか。文部科学省が2012年度から2017年度に実施した「開かれた大学づくりに関する調査」<sup>1)</sup>によると、学生の地域貢献活動<sup>2)</sup>の実施有無について「行っている」と回答した割合は、2013年度の75.9%から年々増加し、2017年度には87.4%に上昇している（表1）。そのうち、取組の実施形態について「課外活動として実施」と回答した割合は、2017年度には87.7%と最も多い。また、「授業の一環で実施」は77.1%であり、2013年度と比較すると16.8%上昇している（表2）。次に、取組の連携状況について2017年度の同報告書によると、大学を設置している自治体との連携の割合が最も高く（87.5%）、他の大学等の教育機関（46.0%）や、NPO法人等の非営利団体（40.9%）、企業（40.4%）等との連携もみられる。

このように、学生が参画する地域貢献の取組は、地域からも期待され、授業の一環で取り組まれる地域貢献プログラムは、益々増えることが予測されている（野上 2019、p.35）。こうしたプログラムは、SL（Service Learning）<sup>3)</sup>としてカリキュラムに位置づけられ<sup>4)</sup>、PBL（Project Based

表1 学生の地域貢献活動の実施有無

	2013	2014	2015	2017
行っている	75.9%	83.6%	86.2%	87.4%
行っていない	24.1%	16.4%	13.8%	12.6%

(出所) 文部科学省「開かれた大学づくりに関する調査」を基に筆者作成

表2 学生の地域貢献活動の実施形態

	2013	2014	2015	2017
授業の一環で実施	60.3%	63.8%	73.1%	77.1%
課外活動として実施	79.4%	79.0%	85.7%	87.7%
その他の形態で実施	27.9%	45.9%	64.1%	65.7%

(出所) 文部科学省「開かれた大学づくりに関する調査」を基に筆者作成

LearningまたはProblem Based Learning) やCBL (Community Based Learning) 等の教育方法として、学部や大学全体で取り組む事例がみられるようになってきている。

### 3. 地域貢献プログラムの特徴

本章では、表3の地域貢献プログラムと公共政策の三つの要素から、地域貢献プログラムは政策問題に取り組む公共政策であることを確認し、教育と公共政策の側面を持つことを示す。

地域貢献プログラムの要素について、まず、一つ目は内容の多くが実世界の問題を扱っている点である。大学全体で地域貢献を進めたCOC事業採択校の取組を例にみると、その分野は歴史や文化、健康、福祉、観光、まちづくり、地域産業等、多岐にわたっている<sup>5</sup>。例えば、商店街組合と学生がまちづくりや健康、福祉イベントを企画し、地元商店街の活性化を図る取組や、自治体が抱える課題に対して、学生の視点から解決策の提案を行う取組、地元食材を活かした商品の企画や開発等、様々である。こうしたプログラムの内容は、大学の学部の分野特性や教職員の専門性と連携先の課題をすり合わせ、取り組む内容が決定される。

公共政策学は、福祉や産業、金融、環境等の個別政策分野のみならず、これらの既存のものを分野横断的に応用する学問である(足立 2009、

pp.12-13)。社会で解決すべき公共的問題と認識された問題は、政策問題と称され、こうした公共的問題(政策問題)を解決するための解決の方向性と具体的手段が公共政策とされる(秋吉ほか 2020、p.4)。このような公共政策学の教育方法において窪田(2020)は、PBLがその中核を担っており、政策現場との関わりを持つことで、知識と体験の統合を行う政策形成の体験そのものであると述べている(窪田 2020、p.52)。こうしたPBLは、実世界で直面する問題の解決を通して、問題解決に関する能力や態度等を身に付ける学習である(溝上 2016、pp.6-9)。

このような公共的問題を扱う地域貢献プログラムは、地域や政策系を専門とした一部の学部や学科に限らず、様々な分野で実世界の問題を扱う内容が含まれている(石井 2012、p.91)。つまり、地域貢献プログラムは公共政策と捉えることができる。

二つ目の要素は、地域貢献プログラムが多様な主体により形成され取り組まれている点である。第2章2節で述べたように、地域貢献プログラムは、取組の連携先として自治体や他の大学等の教育機関、企業、非営利団体等があり、その他にも地域の自治会や町内会、ボランティア組織といった地縁組織等ともプログラムの企画や調整を行い、取り組まれている。

公共政策においても、政策に関わるのは、官僚や政治家等、権限を有するアクターだけでなく、業界団体や労働組合、マスコミ等が挙げら

れる。また、地方自治体では、市民が首長や議員を選出するという間接的な関わりだけでなく、審議会等の場に直接参加し、政策案に対して意見を述べることも珍しくなく、多様なアクターが関与している（秋吉ほか 2020、pp.32-33）。双方とも多様な主体が関わっている。

最後に三つ目の要素は、市民社会の担い手を育成している点である。地域貢献プログラムでは、COC事業を例にみると、学生が地域の課題等に対して認識を深めることが重要な要素とされている。学生が人との関わり合いを通じて、地域の課題への理解や課題解決に向き合う姿勢を養い、地域で活躍できる人材として成長することをねらいとしており、大学が地域を支える人材を育成し、現場の担い手として輩出することが期待されている。

公共政策学においても、日本の民主主義社会で公共的な諸問題を公共的な観点から考察し、問題に対する適切な処方箋を発見し構想する実践能力（政策的思考、ポリシーマインド）を持つ人材を育成することが使命の一つとされている（足立 2009、p.17）。双方とも市民社会の担い手として貢献できる人材の育成を目指している点で共通点がみられる。

以上の三つの要素から、大学の地域貢献プログラムは、様々な学問分野で地域社会の課題として捉えられる政策問題をテーマに、大学（学生や教職員）と地域（自治体や企業、非営利団体、地縁組織等）の多様な主体が協働し、市民社会の担い手の育成を視野に入れて取り組ま

れている。このようなプログラムは、授業の一環として行われているが、学生の学習成果と地域の課題解決に寄与することの双方が期待されている。つまり、地域貢献プログラムは教育と公共政策の側面を持つことが特徴である。

#### 4. 地域貢献プログラムに質改善を目的とした評価の意義

では、このような大学の地域貢献プログラムは、実際にどのような評価が行われているのだろうか。先行研究から評価の現状と課題について取り上げ、現場の取組の質を高める（プログラムをより良くする）質改善を目的とした評価は、どのような効果が生じるのか、評価の意義について述べる。

##### (1) 評価の現状

まず、大学の地域貢献プログラムの評価の現状として二つの特徴が挙げられる。一つ目は、文部科学省による支援策に採択された大学は、資金提供者の意図やねらい（アカウンタビリティ）を重視した評価になりやすい点である。大学の地域貢献は全国的に広がっているが、取組の評価を行い、具体的な評価指標を持つ大学はまだまだ少ない。

こうした背景から木下（2020）は、COC事業の採択校に着目し、採択校には取組に対する評価が義務付けられていることから、具体的な評

表3 地域貢献プログラムと公共政策の要素

	地域貢献プログラム	公共政策
①実世界の問題を扱う	<ul style="list-style-type: none"> <li>扱うテーマは、歴史や文化、健康、福祉、観光、まちづくり、地域産業など多岐にわたる</li> <li>大学の分野特性や教職員の専門性と連携先の課題をすり合わせ、取り組む内容を決定</li> </ul>	<ul style="list-style-type: none"> <li>政策問題を解決するための解決の方向性と具体的な手段</li> <li>公共政策学教育の教育方法では、PBLがその中核を担う</li> <li>PBLは実世界の問題解決に取り組むことを特徴とする</li> </ul>
②多様な主体により形成	<ul style="list-style-type: none"> <li>自治体や教育機関、企業、非営利団体、地域のボランティア組織、地縁組織（団体）等と連携</li> </ul>	<ul style="list-style-type: none"> <li>官僚、政治家、業界団体、労働組合、マスコミ、市民等が関わる</li> </ul>
③市民社会の担い手の育成	<ul style="list-style-type: none"> <li>学生が地域課題への認識を深める</li> <li>地域課題の解決に向けて主体的に行動する人材を育成</li> </ul>	<ul style="list-style-type: none"> <li>政策的思考（政策問題を公共的観点から考察し、問題に対する適切な処方箋を発見し構想する実践能力）を持つ人材の育成</li> </ul>

（出所）筆者作成

価値指標や評価結果をフィードバックする仕組みを持つ大学4校にヒアリング調査を行った。その結果、採択をきっかけに組織内に一定程度の評価や仕組みが構築されたと回答した大学が確認されている。調査を行ったサンプルは少ないが、文部科学省がCOC事業の採択校へ評価を義務付けたことは、大学が地域貢献の取組の評価体制を構築するうえで有効であったのではないかと考察している（木下 2020、p.133）。

一方で、資金提供者から評価を義務付けられるということは、資金提供者の意図を汲み取り提示される一定の条件を満たした評価を行わなければならないことが想定される。例えばCOC事業では、2015年度からCOC+事業に統合され、学生の就職先の創出や地元就職率の向上を推進する取組が対象になった<sup>6</sup>。これにより、採択校は学生の就職を視野に入れた取組を考案することや連携先が限定されること、また、評価の観点は、文部科学省により提示された条件（COC+事業の場合は、地元就職率）を考慮しなければいけない。

二つ目は、地域貢献プログラムの評価は、学生の学習成果を目的としたものが中心であることである。中川・荻野（2020）は、CBLの評価に関する先行研究を概観し、学生の学習成果に関する研究が多く、その代表的な評価方法としてルーブリック評価を挙げている（中川・荻野 2020、p.39）。また、大学教育の現場では教育の質が問われ、学習成果測定を含む教育の内部保証システムの構築が喫緊の課題である。大学基準協会では、教育プログラム評価に教育と地域貢献を結びつけた付加的プログラムを視野に入れて議論し、教育改善と大学の質保証、説明責任を主な評価の目的としている（大学基準協会 2019、pp.8-9）。

## (2) 評価の課題

前節の評価の現状を踏まえ、地域貢献プログラムの評価について、現在、行われている評価と、今後行うことが必要な評価（課題）について図1のように整理した。第3章で述べたように、地域貢献プログラムは、教育と公共政策の側面を持っている。つまり、地域貢献プログラムの

評価については、教育と公共政策の視点から評価を行うことが必要である。本節では、図1で示した学生の学習成果、アカウントビリティ、質改善の三つの観点から評価の現状を整理し、地域貢献プログラムの評価の課題について提示する。

まず、学生の学習成果の評価については、学生がプログラムを通じて何を学んだのか、知識や技術がどの程度身に付いたのか、具体的に何を習得したのか等、授業中のグループ学習の様子や授業後のレポートや報告書、ルーブリック評価等によって学習成果が確認されている。前節でも述べたように、SLやPBL等における学生の学習成果については、先行研究や報告書等が数多くみられる。

次に、アカウントビリティに関する評価については、前節で、資金提供者の意図やねらい（アカウントビリティ）を重視した評価になりやすい現状について述べた。アカウントビリティを目的とした評価は、本来、プログラムの質改善に繋がりうるものであるが、資金提供者の意図や提示された条件を考慮することが必要であること、また、現状としては、資金提供者への報告（報告書の作成）が中心となり、評価結果の報告に留まるケースが多い。評価結果の活用やフィードバック、改善にはなかなか至っていないのが現状である。

最後に、質改善の評価について、例えば、地域貢献プログラムを実施したことによる連携先に生じた効果の有無や効果があった場合、具体的にどのような変化が生じたのか、また、プログラムを実施し問題はないのか等、連携先への影響を考慮した評価を行っている事例は数少ない。大学によっては、プログラムの実施後に連携先の組織や地域コミュニティの代表となるキーパーソンへのインタビュー、アンケート調査等により評価を行っているケースが考えられる。しかし、冒頭でも述べたように、学生の学習成果の評価に関する研究に比べ、地域への影響を考慮した先行研究は数少ないのが現状である（荻野・中川 2021、p.131）。

このような地域貢献プログラムは、地域の経験や知識を高等教育に活かすという教育の視点

は強いが、一方で、地域がどうなることが大学の社会貢献に繋がるのか、地域への影響に対する検証は不十分であることが問題提起されている(村田 2019, p.78 ; 宮崎ほか 2013, p.11)。また、プログラムを通じて、どの程度、課題解決に貢献することができたのか、連携先の視点からプログラムに対する満足度等について、異なる観点から議論される必要性も指摘されている(石井 2012, p.92 ; 山口 2017, p.39)。今後は、地域のために高等教育がどのように貢献することができるのかという視点で研究することも必要である(中川・荻野 2020, pp.40-41)。

このように、地域貢献プログラムの評価の課題については、公共政策の視点からプログラムをより良くするための質改善を図る評価は、学生の学習成果に関する研究に比べ数少なく、検証が不十分であり今後の課題といえる。

では、質改善を目的とした評価とはどのような評価だろうか。例えば、プログラム評価の視点を取り入れ、プログラムの主な実施主体となる大学の教職員が中心となってロジックモデルを作成し、プログラムの協力者である連携先と共有し、双方の目的やねらいをすり合わせ、合

意形成を図ることが重要であると考えられる。この際、プログラムの内容や実施したことによる大学と連携先の成果(アウトカム)を視覚化することで、実施後にプログラムの効果や課題を検証することが可能である。また、ロジックモデルを作成し、連携先と合意形成を図ることは、プログラムの目的やねらいを共有するだけでなく、連携先の期待と大学が提供できるサービス(資源)の食い違いを防ぐことにも繋がる。連携先だけでなく、プログラムの参加主体である学生においても、プログラムを通して身に付けてもらいたい能力や期待する成果を達成することができたのか確認することが可能である。このように、大学と連携先の双方にとってプログラムをより良くするための質改善を図る評価を行うことができると考えられる。

### (3) 質改善の評価を行うことの意義

では、このような質改善を図る評価を行うことで、どのような効果が期待できるのか。本節では、図2に示した大学の代表的な三つの評価を例に挙げ、質改善を目的とした評価を行うことで生じる効果の可能性について述べたい。

まず一つ目は、認証評価である。これは、大学等の教育や研究の質向上と改善を支援することを目的としており、全ての大学、短期大学、高等専門学校は七年に一回、文部科学大臣の認証を受けた評価機関(認証評価機関)から評価を受けることが学校教育法により義務付けられている。例えば、認証評価機関の一つである日本高等教育評価機構では、評価基準として六つの項目を設けている<sup>7</sup>。これらの項目の中には、社会貢献や地域貢献に関する項目はないが、項目以外に各大学が個性や特色として重視している領域については、独自に基準項目を設定することが可能である。近年、こうした地域貢献の取組が増えており、各大学が作成し公開される「自己点検・評価報告書」に独自項目として、記載する大学もいくつかみられる。しかし、内容の多くが取組の実施件数や概要等が中心であり、具体的な取組の効果を検証し記述された評価報告書は見当たらない。

質改善を目的とした評価を行うことにより、

図1 地域貢献プログラムの評価の現状と課題



(出所) 筆者作成

例えば、プログラムを実施したことによる連携先の効果の有無や具体的な内容を検証することで、その結果を「自己点検・評価報告書」に記載し情報公開することで、社会全体にアピールすることができる。これは、大学の独自性をより強化することに繋がり、大学側にもメリットがあると考えられる。

また、冒頭でも述べたように、政府は大学が社会課題を解決する担い手の一つであると認識しており、地域の課題解決に貢献できるような様々な支援策を講じて大学の地域貢献を後押ししている。質改善を目的とした評価では、大学の視点（自己評価）だけでなく、連携先の視点も含まれるため、プログラムを実施したことでのような効果が生み出されたのか、大学が地域に貢献することができた具体的な内容について示すことが可能である。

二つ目は、授業評価である。これは、教育改善を目的としており、文部科学省によると、学生による授業評価の実施状況としては、国公私立全体で756大学（約99%）となっており<sup>8</sup>、ほぼ全ての大学で実施されている。授業後、主にアンケート調査によって学生の意見を集約する方法や、学生により結成された自治組織から意見聴取の機会を設けること等が挙げられる。例えば、授業後に実施されるアンケート調査の主な評価の観点は、学生の授業への出席状況や態度（意欲）、理解度、授業の分かりやすさ、教員の授業に対する熱意や意欲、話し方や教材の適切さ、学習環境（教室の広さや音響等の設置機材）等である<sup>9</sup>。これにより、学生の授業に対する意見や改善点を一定程度は把握することができる。

しかし、地域貢献プログラムの場合、連携先の視点から考えると、プログラムが連携先にどの程度貢献することができたのか、課題解決に寄与することができたのか等、連携先から意見を聴取する機会を設けない限り把握することはできない。

質改善を目的とした評価を行うことで、例えば、あらかじめロジックモデルを作成することにより、連携先の意図したねらいや成果を達成することができたのか、達成できていなければどこに問題があったのか等、今後の課題を検証

することができる。第3章でも述べたように、地域貢献プログラムの特徴として、実世界の問題を扱うことや連携先として多様な主体が関わっていることを取り上げた。特に、大学が提供できるサービス（資源）と連携先が期待することをマッチングし双方のニーズを満たすプログラムの設計や目的の設定は、プログラムを実践する際の課題の一つとして挙げられる（石井 2012、p.92；林ほか 2017、p.113；S.ゲルモンら 2015、p.123）。

つまり、評価を行う際の視点としては、大学側が得られた効果だけでなく、連携先に生じた効果も検証することが必要である。このような質改善を目的とした評価結果を活用することで、大学と連携先の双方の意見を反映させたプログラムづくりに繋げることが可能である。

三つ目は、学習評価（成績評価）である。これは、教員が学生の学習到達度を確認することを目的としており、主にペーパーテストや授業後のレポート、ルーブリック評価等の方法が挙げられる。これらは、教員が問題や評価項目をあらかじめ設定し、教員が期待するねらいや能力を学生自身が身に付けることができたのかを確認する評価である。また、学習評価の評価主体は、教員だけでなく学生自身が自分の学びを振り返る目的で行う自己評価や学生同士の相互評価という形で行われる場合もある。

一方で、連携先に生じた変化（効果や課題等）については、連携先から意見を聴取する機会を設けない限り把握することはできない。

このような連携先の成果を評価することは、前述したプログラムの改善だけでなく、学生の学習面においても非常に重要である。例えば、学生が連携先とともにプログラムに参加したことで、連携先にどのような影響（プラス面やマイナス面）を与えたのか、具体的な効果や課題について振り返り、学生にフィードバックすることができる。連携先に生じた具体的な効果の内容を伝えることは、学生自身がプログラムに取り組んだ達成感に繋がる。問題が生じていれば、その要因は何かを考えるきっかけとなり、対策を検討し今後の課題について振り返ることができる。このように、連携先への影響を検証

することは、学生がより深い学びに繋げるきっかけとなるため、非常に重要な要素であることが先行研究においても示唆されている（中村2013, pp.9-14；S.ゲルモンほか2015, pp.38-40）。

また、第3章2節でも述べたように、質改善を目的とした評価は、学生と連携先の成果（アウトカム）を視覚化したロジックモデルを作成し、双方の視点を踏まえた評価を行うため、市民社会の担い手の育成という観点からプログラムを検証することも可能である。

例えば、学生は地域に関心を持ち、地域が持つ社会資源について理解できたか、地域課題を発見し、課題解決に向けた具体的な提案や実践に向けて主体的に行動できる力を身に付けたか等、教員が意図したねらいが達成されたのか検証される。一方、連携先には実施されたプログラムを通じて、例えば、学生が地域に入ったことにより、変化が生じた、活気が出た、今までにない斬新なアイデアがもたらされた等、どのように地域に貢献することができたのか検証される。このように、学生と連携先の双方の視点から評価を行うことで、市民社会の担い手に寄与するプログラムに至ったのか検討し、見直すことが可能である。

## 5. おわりに

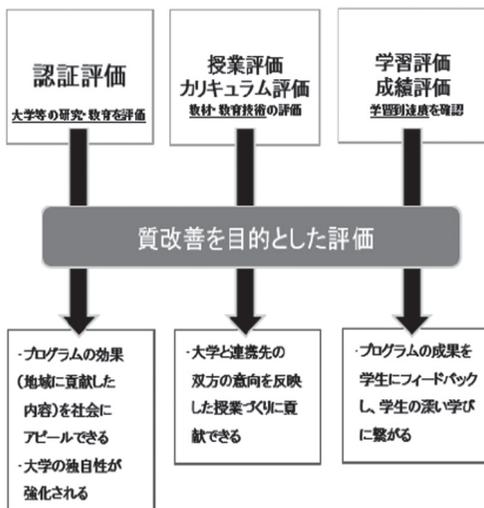
本稿では、授業の一環として行われている大学の地域貢献プログラムの特徴について、地域貢献プログラムと公共政策の三つの要素から整理し、教育と公共政策の側面があることを示した。また、こうした地域貢献プログラムの評価の主な目的は、資金提供者へのアカウンタビリティや学生の学習成果が中心となっており、プログラムをより良くするための質改善の評価の研究は数少なく、検証が不十分であることを課題として提示した。最後に、質改善を図る評価を行うことの意義として、地域貢献プログラムに関わる大学の代表的な三つの評価を例に挙げ、評価を行うことで生じる効果の可能性について述べた。

今後の課題としては、プログラムの質改善を目的とした評価の具体的な評価方法が挙げられる。第4章3節では、大学で行われている代表的な三つの評価を例に挙げたが、その他にも各大学で様々な評価が行われている。特に、地域貢献プログラムを実践する教職員は、プログラムの企画や連携先との調整、学生の引率や指導、実践に向けた準備等、通常業務に加えて目に見えにくい様々な業務を抱えている。また、活動時の費用の捻出や授業の一環で行われることから時間の制約もある。そのため、評価方法を検討する際は、プログラムを実践する教職員の評価疲れや負担感等を考慮しなければいけない。また、評価の負担については、プログラムの実施主体となる教職員だけでなく、協力者となる連携先にとっても同様の検討課題である。連携先がプログラムの評価に参加することで得られるメリットや評価を行うことの必要性についても理解を図る必要がある。ここで取り上げた課題については、今後の研究課題として取り組みたい。

## 謝辞

本論文の原型は、日本評価学会第22回全国大会での報告である。司会を務めていただき、建設的なコメントやご指摘をいただいた。山口大

図2 質改善の評価を行うことで生じる効果



(出所) 筆者作成

学・経済学部の西山慶司先生に深謝する。また、フロアから質問をいただいた方々にも御礼申し上げます。

## 注記

- 1 大学の地域貢献に関する調査は様々行われているが、全国の大学・短期大学を対象にアンケート調査を実施し、いずれの年も回収率が8～9割と高いため本調査を選定した。
- 2 本調査では「学生が、地域が抱える諸問題の発見とその解決に向けて、地域の人々（市民）と協力しながら行う活動」と定義されている。
- 3 学修者の主体的（能動的）な学びを促すアクティブラーニングの一つであり、奉仕（サービス）と学習（ラーニング）を結びつけ、学修者と社会との結びつきを意識した教育的取組のこと。
- 4 高等教育においては、教育方法の質的転換を求めるものとして、中央教育審議会の答申で「サービス・ラーニング科目、NPOに関する専門科目等の開設」（2002年）を進めることが提唱され、この頃からSLを導入し、関連の常設機関を設置する大学もみられるようになった（宮崎ほか2013、p.5）。
- 5 日本学術振興会「平成28年度評価 評価結果一覧」（最終閲覧日：2021年10月10日確認）  
[https://www.jsps.go.jp/j-coc/h28\\_hyoka\\_kekka.html](https://www.jsps.go.jp/j-coc/h28_hyoka_kekka.html)
- 6 日本学術振興会「地（知）の拠点大学による地方創生推進事業（COC+）制度概要」（最終閲覧日：2021年10月15日確認）、  
<https://www.jsps.go.jp/j-coc/gaiyo.html>
- 7 評価基準として、①使命・目的等、②学生、③教育課程、④教員・職員、⑤経営・管理と財務、⑥内部質保証の6つの基準を設定している。この評価基準は、教育を中心とした大学の基本的・共通的な内容で構成されており、基準項目ごとに各大学が満たすことが必要な内容が定められている。
- 8 文部科学省高等教育局・大学振興課大学改革推進室「平成30年度の大学における教育内容等の改革状況について（概要）」（最終閲覧日：2022年2月17日）  
[https://www.mext.go.jp/content/20201005-mxt\\_daigakuc03-000010276\\_1.pdf](https://www.mext.go.jp/content/20201005-mxt_daigakuc03-000010276_1.pdf)：22
- 9 文部科学省「学生による授業評価」（最終閲覧日：2021年10月15日確認）[https://www.mext.go.jp/a\\_menu/](https://www.mext.go.jp/a_menu/)

[koutou/da\\_igaku/04052801/002.htm](https://www.mext.go.jp/a_menu/)

## 参考文献

- 秋元みどり（2020）「大学と地域のパートナーシップのアセスメントをめぐって—大学と市民活動団体との連携事例を通じた検討—」、『日本福祉教育・ボランティア学習学会研究紀要』、(35)：24-38
- 秋吉貴雄・伊藤修一郎・北山俊哉（2020）『公共政策学の基礎〔第3版〕』、有斐閣
- 足立幸男（2009）『公共政策学とは何か』、ミネルヴァ書房
- 石井雅章（2012）「課題解決型教育（PBL）による地域課題解決への貢献」、『地方自治研究』、27（1）：83-96
- 荻野亮吾・中川友理絵（2021）「大学と地域のパートナーシップの質と地域に与える影響の評価方法の検討—高等教育機関における地域と連携した学習に関するレビューから—」、佐賀大学大学院学校教育学研究科紀要、(5)：117-135
- 木下由梨枝（2020）「大学の地域貢献プログラムにおける評価方法のあり方の一考察—「地（知）の拠点整備事業（COC事業）」採択校に着目して—」、『大学評価研究』、(19)：125-135
- 窪田好男（2020）「公共政策学教育におけるケース・メソッドの重要性—PBLとの相違を中心に—」、『公共政策研究』、(20)：49-61
- 公益財団法人大学基準協会（2019）『教育プログラム評価ハンドブック』、公益財団法人大学基準協会
- 公益財団法人日本高等教育評価機構（2021）「大学機関別認証評価 実施大綱」（最終閲覧日：2022年2月14日）<https://www.jihe.or.jp/achievement/col lege/pdf/jisshitaiko200401.pdf>
- 田中耕治編（2021）『よくわかる教育評価 第3版』、ミネルヴァ書房
- 田中優（2018）『学生・教職員・自治体職員・地域住民のための地域連携PBLの実践』、ナカニシヤ出版
- 地域・大学協働研究会（2014）『地域・大学協働実践法 地域と大学の新しい関係構築に向けて』、悠光堂
- 中川友理絵・荻野亮吾（2020）「日本の高等教育における地域と連携した学習（Community-Based Learning）の研究動向」、『日本地域政策研究』、(25)：34-43

- 中村みどり(2013)「地域社会との協働性を高めるプログラムづくりに向けて—大学教育に求められる地域への視点—」、『キリスト教教育研究』、(30)：1-24
- 野上真(2019)「わが国の大学における地域連携の動向と展望」、志學館大学研究紀要、(40)：27-44
- 野澤一博(2016)「大学の地域連携の活動領域と課題」、『産学連携学』、13(1)：1-8
- 林透・眞鍋和博・猪股歳之・岩瀬峰代・西村君平・山下貴弘(2017)「ラウンドテーブル報告 地域連携学習(コミュニティー・ベースト・ラーニング)の設計・運営・評価とその担い手のあり方について考える」、『大学教育学会誌』、39(2)：110-114
- 溝上真一・成田秀夫(2016)『アクティブラーニングシリーズ 第2巻 アクティブラーニングとしてのPBLと探求的な学習』、東信堂
- 宮崎猛・伊藤章・眞所佳代(2013)「サービス・ラーニングにおけるコミュニティインパクト(貢献活動の影響)への捉え—日米の高等教育機関への調査を基に—」、『学士課程教育機構研究誌』、(2)：5-23
- 村田和子(2019)「『地域づくり』と大学生涯学習：『大学地域連携』政策の批判的検討」、日本社会教育学会〔編集〕、『地域づくりと社会教育的価値の創造』、東洋館出版社、66-79
- 文部科学省「開かれた大学づくりに関する調査」(最終閲覧日：2021年9月8日) [https://www.mext.go.jp/a\\_menu/ikusei/daigaku/1288601.htm](https://www.mext.go.jp/a_menu/ikusei/daigaku/1288601.htm)
- 山口泰史(2017)「わが国におけるPBL研究の動向：大学教育での実践を中心に」、『日本地域政策研究』、(19)：34-41
- 安田節之(2011)『プログラム評価 対人・コミュニティ援助の質を高めるために』、新曜社
- 山谷清志〔監修〕・源由理子・大島巖〔編著〕(2020)『プログラム評価ハンドブック—社会課題解決に向けた評価方法の基礎・応用—』、晃洋書房
- S. ゲルモン・B.A. ホランド・A. ドリスコル・A. スプリング・S. ケリガン〔著者〕、山田一隆〔監訳〕、市川享子・齋藤百合子・福井里江・村上徹也・中原美香〔訳〕(2015)『社会参画する大学と市民学習 アセスメントの原理と技法』、学文社

(2022.6.30 受理)

## **Significance of conducting evaluations for the purpose of quality improvement in university community contribution programs**

Yurie Kishimoto

Niimi University  
yk6710@niimi-u.ac.jp

### **Abstract**

In recent years, university contributions to local communities have spread nationwide with various government support measures. It is important to link these efforts to education and research to achieve synergistic effects, and the number of community contribution programs conducted as part of classes has been increasing every year and is expected to continue to increase in the future.

This paper focuses on programs that link university education with community contributions and are offered as part of the coursework. Although these programs are characterized by aspects of education and public policy, we present as an issue that evaluations aimed at improving the quality of the programs from a public policy perspective are few in number and insufficiently verified compared to studies on student learning effectiveness. We will also discuss the concrete effects of quality improvement evaluation, and the significance of evaluation for the purpose of quality improvement, using three typical evaluations conducted at universities as examples: accreditation evaluation, class evaluation, and learning evaluation (grade evaluation).

### **Keywords**

Community Contribution Program, Public Policy, quality improvement, program evaluation, Logic Model



## 日本評価学会 第23回全国大会のご案内

企画委員長 齊藤 貴浩 (大阪大学)  
第23回全国大会実行委員長 贅川 信幸 (日本社会事業大学)

日本評価学会第23回全国大会については、昨今の新型コロナウイルス感染症(COVID-19)の影響等に鑑み、オンラインにて開催いたします。詳細は学会メーリングリストおよびホームページにてご案内いたします。奮ってご参加ください。

### 記

1. 日程：2022年12月10日(土)、12月11日(日)
2. 方法：オンラインにて開催(前回と同様にZoomを用います)

以上



『日本評価研究』編集委員会  
Editorial Board

編集委員長  
Editor-in-chief

西野 桂子(関西学院大学)  
Keiko NISHINO

常任編集委員  
Standing Editors

大島 巖(東北福祉大学)  
Iwao OSHIMA

牟田 博光(東京工業大学)  
Hiromitsu MUTA

山谷 清志(同志社大学)  
Kiyoshi YAMAYA

米原 あき(東洋大学)  
Aki YONEHARA

編集委員  
Editors

窪田 好男(京都府立大学)  
Yoshio KUBOTA

贄川 信幸(日本社会事業大学)  
Nobuyuki NIEKAWA

橋本 昭彦(国立教育政策研究所)  
Akihiko HASHIMOTO

林 薫(文教大学)  
Kaoru HAYASHI

村上 裕一(北海道大学)  
Yuichi MURAKAMI

特定非営利活動法人日本評価学会  
編集委員会事務局  
E-mail: [journal@evaluationjp.org](mailto:journal@evaluationjp.org)

～投稿案内～

日本評価学会では、「日本評価研究」掲載のための投稿原稿を募集しております。投稿の締め切りは9月末日（翌年3月刊行）及び3月末日（9月刊行）です。

ご興味をお持ちの方は投稿規定・執筆要領・査読要領、並びに原稿見本をご参照のうえ、奮ってご投稿ください。<http://evaluationjp.org/activity/journal.html>

投稿先：特定非営利活動法人日本評価学会 編集委員会事務局  
E-mail：journal@evaluationjp.org

『日本評価研究』第22巻第2号  
2022年9月30日

編集・発行 特定非営利活動法人日本評価学会 編集委員会  
〒113-0034 東京都文京区湯島3-31-1  
中川ビル5階  
一般財団法人行政管理研究センター内

印刷 株式会社 研恒社

©日本評価学会

本誌に掲載されたすべての内容は、日本評価学会の許可なく転載・複写はできません。

# Japanese Journal of Evaluation Studies

Vol. 22, No. 2, September 2022

## CONTENTS

### Special Issue: Formative evaluation for “innovation programs”

- Editor’s Note: For the development of effective formative evaluation approaches  
Iwao Oshima
- Development of a formative evaluation approach that can foster human service innovations into Evidence-Based Practices: Evolvability of the Practitioner and peer Based Empowerment Evaluation using CD-TEP: the PBEE/CD-TEP approach  
Iwao Oshima, Kenta Shindo, Yuriko Minamoto
- Trend and Dynamics of Participatory Formative Evaluation in the Field of Education  
Aki Yonehara
- From Social Impact Measurement to Evaluation for Social Impact: Evolution of the Concept and Practice towards Evaluation for Management Support  
Katsuji Imata
- UK initiatives to develop innovations in human services into Evidence-Based Programme  
- Focus on the What Works Centres in the child welfare -  
Naoyuki Ieko
- Development and Implementation of an Evaluation Guide for Developing Field Practices into an Effective Model - A Case Study Using the CD-TEP Approach -  
Junko Shimizu, Kenta Shindo
- The Significance and Role of the EBP (Evidence-Based Practices) Technical Assistance Center, the Foundation for Practitioner - Based Empowerment Evaluation -Focusing on the Evaluation Capacity Building -  
Kenta Shindo, Iwao Oshima
- Article
- Promote the Use of Local Government Accounting Information as Open Data  
open data for municipal accounting information: Focusing on the classification of expense by COFOG  
Yuichi Ubukata, Toru Satou, Hiroichi Kawashima
- Institutionalizing Systematic Evaluation to Ensure “Better Regulation”  
- Focusing on Regulatory Oversight Body and Experiment Evaluations -  
Yuichi Murakami
- Research Note
- The need of inductive information to improve the quality of indicators in evaluation by performance measurement  
Hazuki Ikeda
- An Evaluation of Air Pollution Control in Ulaanbaatar  
Batkhuuyag Undrakh
- Significance of conducting evaluations for the purpose of quality improvement in university community contribution programs  
Yurie Kishimoto